

**ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМий ДАРАЖАЛАР БЕРУВЧИ
DSc.13/30.12.2019.Т.07.01 РАҚАМЛИ ИЛМий КЕНГАШ**

ГУЛИСТОН ДАВЛАТ УНИВЕРСИТЕТИ

ТИШЛИКОВ СУЛТОНЖОН АБДУРАИМОВИЧ

**ВИЛОЯТ БОШҚАРУВИДА «ЭЛЕКТРОН ҲУКУМАТ» ТИЗИМИ
ҲУЖЖАТЛАРИГА ИШЛОВ БЕРИШ ИШОНЧЛИЛИГИНИ ОШИРИШ
УСУЛЛАРИ ВА АЛГОРИТМЛАРИ**

05.01.02 - Тизимли таҳлил, бошқарув ва ахборотни қайта ишлаш

**ТЕХНИКА ФАНЛАР БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Тошкент–2021

**Техника фанлари бўйича фалсафа доктори (PhD) диссертацияси
автореферати мундарижаси**

**Оглавление автореферата диссертации
доктора философии (PhD) по техническим наукам**

**Contents of dissertation abstract of doctor of philosophy (PhD)
on technical sciences**

Тишликов Султонжон Абдураимович

Вилоят бошқарувида «электрон ҳукумат» тизими ҳужжатларига ишлов бериш ишончлилигини ошириш усуллари ва алгоритмлари 5

Тишликов Султонжон Абдураимович

Методы и алгоритмы повышения достоверности обработки документов систем «электронного правительства» областного управления 19

Tishlikov Sultonjon Abduraimovich

Methods and algorithms to increase the reliability of document processing systems of «electronic government» of the regional administration 37

Эълон қилинган ишлар рўйхати

Список опубликованных работ

List of published works 41

**ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМий ДАРАЖАЛАР БЕРУВЧИ
DSc.13/30.12.2019.Т.07.01 РАҚАМЛИ ИЛМий КЕНГАШ**

ГУЛИСТОН ДАВЛАТ УНИВЕРСИТЕТИ

ТИШЛИКОВ СУЛТОНЖОН АБДУРАИМОВИЧ

**ВИЛОЯТ БОШҚАРУВИДА «ЭЛЕКТРОН ҲУКУМАТ»ТИЗИМИ
ҲУЖЖАТЛАРИГА ИШЛОВ БЕРИШ ИШОНЧЛИЛИГИНИ ОШИРИШ
УСУЛЛАРИ ВА АЛГОРИТМЛАРИ**

05.01.02- Тизимли таҳлил, бошқарув ва ахборотни қайта ишлаш

**ТЕХНИКА ФАНЛАР БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Тошкент–2021

Техника фанлари бўйича фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси хузуридаги Олий аттестация комиссиясида В2021.4.PhD/T1296 -рақам билан рўйхатга олинган.

Диссертация Гулистон давлат университетига бажарилган.

Диссертация автореферати уч тилда (ўзбек, рус, инглиз (резюме) Илмий кенгаш веб-саҳифасида (www.tuit.uz) ва «ZiyoNet» Ахборот таълим порталида (www.ziyounet.uz) жойлаштирилган.

Илмий раҳбар:

Жуманов Исроил Ибрагимович
техника фанлари доктори, профессор

Расмий оппонентлар:

Бабомурадов Озод Жураевич
техника фанлари доктори, катта илмий ходим
Сейтов Айбек Жумабаевич
техника фанлари бўйича фалсафа доктори (PhD),
катта илмий ходим

Етакчи ташкилот:

Тошкент давлат транспорт университети

Диссертация ҳимояси Тошкент ахборот технологиялари университети хузуридаги DSc.13/30.12.2019.T.07.01 рақамли Илмий кенгашнинг 2022 йил «14» сентябр соат 16⁰⁰ даги мажлисида бўлиб ўтади. (Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43, факс: (99871) 238-65-52, e-mail: tuit@tuit.uz Тошкент ахборот технологиялари университети).

Диссертация билан Тошкент ахборот технологиялари университети Ахборот-ресурс марказида танишиш мумкин (138- рақам билан рўйхатга олинган). (Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-65-44).

Диссертация автореферати 2021 йил «29» декабр да тарқатилди.
(2021 йил «20» декабр даги 44- рақамли реестр баённомаси.)



Handwritten signature of R. X. Xamdamov

Р.Х.Ҳамдамов
Илмий даражалар берувчи илмий кенгаш раиси,
техника фанлари доктори, профессор

Handwritten signature of F. M. Nuraliyev

Ф.М.Нуралиев
Илмий даражалар берувчи илмий кенгаш илмий
котиби, техника фанлари доктори, доцент

Handwritten signature of A. V. Kobulov

А.В.Кобулов
Илмий даражалар берувчи илмий кенгаш қошидаги
илмий семинар раиси, техника фанлари доктори,
профессор

КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳонда ахборотни қайта ишлаш, тармоқланган WEB орқали бошқарув тармоқлари ва иш жойларини, автоматлаштириш технологияларини ишлаб чиқариш жараёнлари бошқарувини комплекс автоматлаштириш, рақамлаштириш усул ва алгоритмлари, шу жумладан, электрон ҳужжат алмашув тизимлари (ЭҲАТ) билан интеграциялаш йўналишида уларни такомиллаштириш мақсадга мувофиқ эканлиги таъкидланмоқда. Ҳозирги вақтда ҳужжатли ахборотларни қайта ишлашни оптималлаштириш бўйича бажарилган илмий тадқиқотлар натижалари, жуда кўп сонли халқаро компаниялар нашрларида акслантирилган, шунингдек, Gartner, International Data Corporation, Wobot, Forrester, CNews Analytics, ЮНИПРАВЭКС, Интерфейс IT, Интеллект Телеком ва ҳар хил илмий манбаларда ишлаб чиқариш – технологик мажмуалари бошқарувида жуда ҳам зарур бошқарув ахборот тизимлари, маъмурий орган ЭҲАТ ни қарор қабул қилиш учун улар таркибга киритиш ҳамда такомиллаштириш вазифаларини бажариш эҳтиёжи жуда ҳам катталиги асосланади.

Жаҳонда ахборот ишончилигини ошириш бўйича жадал изланишлар олиб борилмоқда ва хорижий мамлакатлар олимлари томонидан муайян муваффақиятларга эришилмоқда. Уларни изланишлари натижаларининг афзаллиги шундаки, улар хатоларни аниқлаш ва тузатиш механизмларини куриш билан электрон ҳужжатлар (ЭХ) матнларининг ишончилигини, хавфсизлигини, яхлитлигини таъминлаш учун илмий ва услубий асослар, усуллар, моделлар ва воситаларни шакллантириш имконини беради. Бу соҳада, АҚШ, Россия Федерацияси, Англия, Франция, Хитой, Ҳиндистон, Германия, Австрия, Швеция, Норвегия, Финландия, Туркия, Исроил, Малазия, Канада, Испания, Жанубий Корея, Бразилия ва бошқа давлатларда шу масалалардаги илмий тадқиқотларга алоҳида этибор берилмоқда.

Ўзбекистонда фан, таълим, рақамли иқтисодиёт, инновацион ғоя ва технологиялар янги усулларни кенг қўламда қўллаш ва ривожлантириш ҳамда замонавий ахборотга ишлов бериш технологияларини турли предмет соҳаларда тадбиқ этиш, айниқса, инсон фактори таъсирини сусайтириш учун самарали воситалар сифатида ҳужжатлар маълумотларини киритиш, узатиш, хотиралаш, қайта ишлаш жараёнларида қўлланишлари учун ахборот ишончилигини ошириш технологияларини ишлаб чиқиш долзарб илмий тадқиқот сифатида ажратилиши ва янгиликни ташкил қилган муаммо масалалари ечимларини олиш назарда тутилади. Муаммо хусусияти шундан иборатки, Республикамизда барча иш ҳужжатлари ўзбек тилида юритилади. Фуқароларнинг ҳаммаси ҳам ўзбек тилида матн яратиш қоидаларини билишмайди ва керакли кўникмаларга эга эмаслар, натижада, ЭҲ матнлари таркибида имловий (орфографик) хатолар содир бўлади. Бу эса ўз навбатида бошқарув тизимларида мавжуд шароитда, ахборот ишончилиги кўрсаткичи паст, қиймати эса фойдаланувчи, инсон – оператор, сканерлаш ва таниб олиш

воситалари хатоликлари, алоқа каналида «помеха»лар таъсири ҳамда электрон жиҳозлар носозлиги туфайли эканлигига асосланган. Ахборот ишончилигини оширишга йўналтирилган мазкур тадқиқот, дастлабки маълумот танқислиги, ностационарлик, параметрларда ноаниқлик мавжуд шартларда ЭҲАТ мобиллиги, самарадорлигини оширишга қодир бўлиши, шу туфайли, ахборотни узатиш ва қайта ишлаш жараёнларида ЭҲ ларда етарли даражада, кам харажатларда ахборот ишончилигини ошириш технологияларини яратишда жуда долзарб ва эҳтиёжли ҳисобланади.

Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси», 2018 йил 19 февралдаги ПФ-5349-сонли «Ахборот технологиялари ва коммуникациялари соҳасини янада такомиллаштириш чора-тадбирлари» Фармонлари, 2017 йил 29 августдаги «Ахборот-коммуникация технологиялари соҳасида лойиҳа бошқаруви тизимини янада яхшилашга доир чора-тадбирлар тўғрисида»ги ПҚ-3245-сонли, 2020 йил 6 октябрдаги «Ахборот технологиялари соҳасида таълим тизимини янада такомиллаштириш, илмий тадқиқотларни ривожлантириш ва уларни ИТ-индустрия билан интеграция қилиш чора-тадбирлари тўғрисида»ги ПҚ-4851-сонли Қарорлари ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишга ушбу диссертация тадқиқоти муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Мазкур тадқиқот республика фан ва технологиялар ривожланишининг IV. «Ахборотлаштириш ва ахборот-коммуникация технологияларини ривожлантириш» устувор йўналиши доирасида бажарилган.

Муаммонинг ўрганилганлик даражаси. Илмий адабиётларини ўрганиш шуни кўрсатдики, сўнгги 10-15 йилларда ахборот ишончилигини ошириш бўйича интенсив тадқиқотлар олиб борилмоқда. Матнларни узатиш ва ишлов беришдаги ишончилиқни ошириш услубияти ва морфологик таҳлил қилиш тизимларини яратишга қаратилган тадқиқотлар Г.Г.Белоногов, С.А.Старостин, М.Г. Мальковский, Кемал Офлазер, Ю. Апресян, Киммо Коскеннилар томонидан асосланган. Шунингдек, Ch.W.Bachman, W.W.Peterson, D.Ritchi, E.F.Codd, C.J.Date, Alan Turing, E.N.Gilbert, R.C.Bose, Danish Jamil, Peter Pin-Shan Chen, Чжо Зо Е, В.Б.Кудрявцев, Ю.В. Семенов, В.В. Трофимов, М.В. Ларин, Г.Ю Максимович, Ю.В. Гуляев каби олимлар ишларида ҳам бошқарув тизимларида ахборотни қайта ишлаш назарияси ва амалиётини ривожлантиришга катта эътибор берилган.

Мамлакатимизда ахборотга оптимал ишлов бериш технологиялари назарияси ривожига академик-профессорлар, Д.А.Абдуллаев, Ф.Б.Абуталиев, М.М.Камилов, Т.Ф.Бекмуратов, Х.З.Игамбердиев, Ш.Х.Фазилов, И.И.Жуманов, Р.Х.Ҳамдамов, М.А.Раҳматуллаев, А.А. Саидов А.Р. Ахатовлар салмоқли ҳисса қўшишган. Тадқиқот натижалари шуни кўрсатадики, маълумотга танқислик, параметрли ноаниқлик, катта

ишончсизлик мавжуд бўлган бошқарув тизимлари шароитида маълумотлар хоссаларидан фойдаланувчи механизмлар асосида ахборот ишончилигини кафолатловчи усул, модел, технологиялар яратилиши етарли даражада ўрганилмаган. Аммо, муаммо ечимида, кўп босқичли морфологик ва таркибланишган n - грамм моделлар асосида ЭХ ишончилигини назорат қилиш, имловий ва бошқа турдаги хатоларни аниқлаш ва таҳрир қилиш механизмларини тадқиқ қилиш ва ишлаб чиқиш жиддий аҳамиятга эга эканлиги таъкидланади. Матн ишончилигини луғатсиз ва луғат асосида назорат қилиш, ишончилигини сунъий, статистик, семантик, таркибий–технологик информатсион ортиқчалик бўйича ошириш ҳамда матндаги бир-, икки-, уч-, n - каррали хатоларни кам ўлчовли луғат бўйича тузатиш, шу билан бирга, информация назоратини луғатсиз амалга оширувчи чизиқли, текислик бўйича 8, 9, 10, 11 ва q модуллар йиғиндиларидан фойдаланиш, ортиқчаликка эга бўлган кодлаш механизмларини қўллаш зарурлиги асосланган.

Диссертация тадқиқоти бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация иши устувор йўналишда, Самарқанд давлат университети ва Гулистон давлат университетининг ИТД-17–«Ахборот ва телекоммуникация технологияларини ривожлатириш ва қўлланишини таъминловчи замонавий ахборот тизимлари, тафаккурли бошқариш ва ўргатиш воситалари, илмий-техник маълумот базалари ва дастурий воситаларини ишлаб чиқиш» (2009-2011) ҳамда ТИД-5- «Ахборот технологиялари, телекоммуникация тармоқлари, аппаратли-дастурий воситалар, тафаккурли бошқариш, ўргатиш, жамиятни ахборотлаштириш даражасини оширишга йўналтирилган усул ва тизимларни яратиш» (2012-2014) билан боғланган ИТИ режалари асосида бажарилган.

Тадқиқотнинг мақсади вилоят бошқаруви маъмурий органлари «электрон ҳукумат» тизимларида ҳужжатлар ахбороти ишончилигини ошириш усул, модел, алгоритмлари ва дастурий воситаларини ишлаб чиқиш.

Тадқиқот вазифалари:

ўзбек тилидаги матнларни кўп босқичли морфологик ва таркибланишган n -грамм таҳлил моделлар бўйича луғатли ҳамда луғатсиз усуллар ёрдамида назорат қилиш ва ахборот ишончилигини оширишнинг услубий асослари, модел ва алгоритмларини ишлаб чиқиш;

чизиқли, текислик бўйича, модулли йиғиндилар ҳосил қилиш ва ортиқчаликка эга кодлаштириш механизмлари асосида матн таркибидаги хатоларни аниқлаш ва тузатиш модел ва алгоритмларини ишлаб чиқиш;

ахборот ишончилигини ошириш алгоритмлари параметрларини оптималлаштириш, уларнинг самарадорлигини хатоларни аниқланмаслик эҳтимоли, ахборот назорати мураккаблиги ва ҳаражат нархи мезонлари асосида тадқиқ қилиш;

ахборот ишончилигини оширувчи дастурий мажмуанинг функционал модулларини ишлаб чиқиш ҳамда вилоят бошқаруви маъмурий органлари «электрон ҳукумат» ЭХАТларига тадбиқ қилиш.

Тадқиқот объекти корхона, ташкилот, муассаса, маъмурий бошқарув органлари фаолиятидаги иш юритиш ҳужжатлари, ЭҲАТ.

Тадқиқотнинг предмети хатоларни аниқлаш ва таҳрир қилиш воситалари, чизикли, текислик бўйича модулли йиғиндиларни ҳосил қилиш, матнни ортиқчаликка эга кодлаш, кўп босқичли морфологик ва *n*-граммли таҳлил қилиш механизми, модел, алгоритмлари.

Тадқиқотнинг усуллари ахборотни узатиш ва қайта ишлаш, алгоритмлаштириш, математик дастурлаш, юқори босқичли алгоритмик тиллардан фойдаланиш, қиёфаларни қидириш ва таниш назариялари.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

ахборот ишончилигини оширишни сунъий, табиий, семантик ва таркибий-технологик ахборотлар ортиқчаликларидан фойдаланишга асосланган конструктив ёндашув, модел ва алгоритмлар таклиф этилган;

чизикли, текислик, модулли йиғиндилар ва ортиқчаликка эга кодлаш ҳамда кўп босқичли морфологик ва таркибланишган *n*-грамм таҳлилларга асосланган механизмлар ишлаб чиқилган;

объектни эвристик қидиришда барча вариантларни танлаш, куйдириш, чегаралаш принципларига, Марков занжири бўйича стохастик моделлаштириш механизмлари умумлашувига асосланган моделлар бўйича кўп қаррали хатоларни аниқлаш ва тузатиш механизмлари ишлаб чиқилган ва алгоритмлар такомиллаштирилган;

ахборот ишончилигини оширишни ишончилик, назорат қилиш мураккаблиги ва ҳаражат нарх қиймати мезонлари бўйича алгоритмлар самарадорлиги тадқиқ қилинган ҳамда уларнинг параметрлари оптималлаштирилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

ахборот ишончилигини оширувчи дастурий мажмуанинг функционал модулларини ишлаб чиқилган ҳамда вилоят бошқаруви маъмурий органлари «электрон ҳукумат» ЭҲАТларига тадбиқ қилинган;

n-грамм частотали луғатларга асосланган матн таркибида кўп қаррали хатоларни аниқлаш ва таҳрир қилиш алгоритмлари параметрларини оптималлаштирувчи усул яратилган ва қўлланилган;

«Сфинкс - 4 « ва «CUDA» параллел ҳисоблаш технологиялари АДПи талабига хос бўлган матн орфографиясини назорат қилиш типик воситалари асосида ДМ маъмурий органлар ЭҲАТи мисолида синовдан ўтказилган.

Тадқиқот натижаларининг ишончилиги ҳамда тўғрилиги ишлаб чиқилган алгоритмлар, дастурий мажмуа, функционал модуллар яроқчилиги, назарий ва амалий натижаларнинг ўзаро мослиги, ахборот ишончилигини ошириш масалаларининг реал шароитларда ечилганлиги ҳамда жорийлаштирилганлиги натижалари қиёсий таҳлили асосида изоҳланган.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти ахборот ишончилигини ошириш алгоритмлари параметрларини мақбуллаштириш, уларнинг самарадорлигида ишончилик, назорат қилиш мураккаблиги ва ҳаражат қиймати мезонлари

бўйича тадқиқ қилинганлиги ҳамда барча вариантларни танлаш, куйдириш, чегараланиш, стохастик моделлаштириш бўйича эвристик қидириш, луғатли ва луғатсиз назорат қилиш, луғат ўлчамини мақбуллаштириш механизмлари бўйича бир, икки, уч, n - карра хатоларни аниқлаш ва тузатиш алгоритмлари яратилганлиги билан изоҳланади:

Тадқиқот натижаларининг амалий аҳамияти n - каррали, аралаш алфавитли хатолар, матн орфографияси назорати ва таҳририни амалга оширувчи механизмлар, функционал модуллари ҳамда ахборот ишончилигини оширувчи дастурий мажмуа яратилганлиги билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Вилоятнинг қуйи “электрон ҳукумат” тизимларида иш юритиш ҳужжатлари маълумотларини қайта ишлашни оптималлаштириш ва узатиш ишончилигини ошириш борасида ишлаб чиқилган моделлар, алгоритмлар ва дастурий воситалар натижаларини амалиётга татбиқ этиш асосида:

ахборотларга ишлов бериш ишончилигини оширувчи механизмлар бўйича ишлаб чиқилган дастурий мажмуа, воситалар, тавсиялар «ЎЗБЕКТЕЛЕКОМ» Сирдарё филиали «Сирдарё Телеком» жорий қилинган. (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2021 йил 8 январдаги №33-8/136-сонли маълумотномаси); Натижада, ахборотни назорат қилиш ва ишончилиқни ошириш механизмлари, анъанавий визуал назорат қилиш технологияси кўрсаткичларига нисбатан, маълумот ишончилигини икки-уч тартибга ошириш, мураккаблик ва ҳаражат нархи коэффицентлари қийматларини 5 - 6 маротобагача камайтириш имконини берган; иш юритиш ҳужжатлари маълумотларини қайта ишлашни оптималлаштириш ва узатиш ишончилигини ошириш борасида ишлаб чиқилган моделлар, алгоритмлар ва дастурий воситалар Сирдарё вилояти бандлик Бош бошқармаси фаолиятида жорий қилинган. (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2021 йил 8 январдаги №33-8/136-сонли маълумотномаси) Натижада, анъанавий визуал назорат қилиш механизмларига нисбатан ахборот ишончилигини 2-3 тартибда ошириши ҳамда илмий ишланмалар самарадорлиги ташкилий-маъмурий ҳужжатларни бажарилиши, таҳлили, қайта ишланишига кетадиган моддий ҳаражатлар миқдорини тахминан 15% га камайтириш имконини берган; ахборот ишончилигини ошириш алгоритмлари бўйича ишлаб чиқилган дастурий мажмуа «АПК Оқ олтин» УК фаолиятида жорий этилган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2021 йил 8 январдаги №33-8/136-сонли маълумотномаси). Натижада жорий этилган дастурий мажмуа матнлардаги имловий хатолар сони 92% гача аниқлаш ҳамда турли икки каррали, ўхшаш, транспозитив хатоларни тузатиш, мураккаблик ва ҳаражат нархи коэффицентлари қийматларини 5-6 маротобага камайтириш имкониятларини берган.

Тадқиқот натижаларининг эълон қилинганлиги апробацияси. Тадқиқот натижалари 11 та халқаро ва 2 та республика анжуманларида муҳокамадан ўтказилган.

Тадқиқот натижаларининг эълон қилинганлиги. Диссертация мавзуси бўйича 32 та илмий иш, жумладан, Ўзбекистон Республикаси Олий аттестацияси комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш тавсия этилган илмий нашрларда 10 та илмий мақола, 1 та чет эл ва 9 та республика журналларида нашр этилган, 4 та ЭҲМ дастурларини рўйхатдан ўтказганлик тўғрисидаги гувоҳномалар олинди.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш, тўртта боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловалардан иборат. Диссертация ҳажми 112 бетни ташкил этади.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Кириш қисмида мавзунинг долзарблиги ва зарурияти асосланган, тадқиқот мақсади, вазифалари, объекти ва предметлари тавсифланган, Республика фан ва технологиялари тараққиётининг устувор йўналишларига мослиги кўрсатилган, илмий янгилиги ва амалий натижалари ҳамда уларнинг назарий, амалий моҳияти талқин қилинган, тадқиқот натижаларини жорий қилиш ҳолати, нашр этилган ишлар бўйича маълумотлар келтирилган.

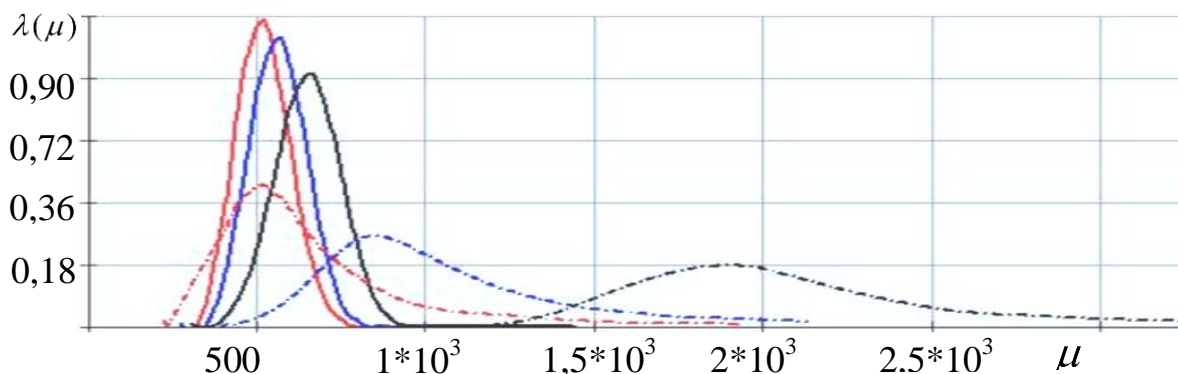
Диссертациянинг «**Бошқариш тизимларида сўз шакллари луғати билан ахборотлар ишончлилигини ошириш механизмлари**» номли биринчи бобида сунъий, статистик, семантик, таркибий-технологик инфор­мацион ор­тик­чалик принциплари, кўп босқичли морфологик ва таркиб­лашган *n*-грамм таҳлил моделлари, чизикли, текислик бўйича модулли йиғиндилар ҳосил қилиш, матнни ор­тик­чаликка эга кодлаш механизмларига асосланган конструктив ёндашув, модел ва алгоритмлар ишлаб чиқилган. Бошқарув маъмурий органлар ахборот-ресурс тармоқлари, ЭҲАТ технологиялари, дастурий сервисларини яратиш масалалари алоҳида ажратиб кўрсатилган. Жумладан, матн ишончлилигини оширишни кўп босқичли морфологик ва *n*-грамм таҳлил моделлари умумлашувида амалга ошириш самарадорлиги таъкидланган.

Қўшимча, тушиб қолган, қўшни ҳарфлар ўрин алмашиши сингари матн хатолари статистикаси тадқиқ қилинган. Ҳар бир сўз шакли учун қўшимча информация, лингвистик билимларни тақдим қилувчи механизмлар, грамматик, синтактик ва морфологик таҳлилларни амалга оширувчи ҳамда *n*-каррали хатоларни аниқлаш ва бартараф қилувчи воситалар умумлашувидаги ахборот ишончлилигини ошириш модел ва алгоритмлари яратилган. Куйдириш, чегаралаш, стахостик моделлаштириш бўйича эвристик қидириш механизмлари, оммалашган кўринишдаги бир, икки каррали транспозитив, ўхшаш, йўқотилган, қўшимча пайдо бўлган ва юқори каррали хатоларни тезкор аниқлаш ва тузатиш механизмларига эга ахборот ишончлилигини ошириш алгоритмлар таклиф этилган. Анъанавий, ўзбек тили матни «spell-checker»и бўйича имловий хатоларни назорат қилиш учун

қўлланиладиган сўзлар учун умумий ва частотали луғатлари ўлчамлари оптималлаштириш услубияти ишлаб чиқилган ҳамда бир, икки каррали, транспозитив, комбинатив кўринишдаги хатоларни назорат ва тузатиш механизмлари умумлашувида алгоритмлар ишлаб чиқилган. Сўз шакли априор фазосини қисиш, тартибли, куйдириш, чеклаш асосида эвристик қидириш, билим базаси (ББ) ва маълумот хоссаларидан фойдаланиш механизмлари бўйича яратилган алгоритмлар такомиллаштирилган ва самарадорлиги тадқиқ қилинган. Луғатсиз алгоритмлар устунлиги хатолар статистикаси, частотали таснифлари, ЭҲ шаблонлари, хато содир бўлишини тавсифловчи моделлари бўйича механизмлар ишлаб чиқилган.

Матнда монограммли ва комбинацияли хатоларни параллел назорат қилиш ва тузатишга қодир адаптив алгоритмлар ишлаб чиқилган. Гиперсемантик қидирув тармоғи, базавий фреймлар жамламаси, қидириш объектлари сонини, қидирув вақтини, ўзгарувчилар қиймати интервалини, мақсад функциясига тўғри ва нотўғри мурожат сонларини, ташқи таъсир интенсивлигини мувофиқлаштирувчи механизмлар ЭҲ матнни орфографиясини назорат ва таҳрир қилувчи умумлашган алгоритмда қўлланилган. Ишлаб чиқилган алгоритм 2000 та фрейм жамламасига эга бўлган 12 та ҳужжат асосида тестлаштирилган. Тадқиқот, алгоритмнинг ахборот ишончилигини назорат қилиш вақти матн орфографиясини визуал назорат қилувчи технологияга нисбатан 6 маротаба кам бўлишини кўрсатди.

Градиент усул бўйича, глобал ва локал экстремумларни қидириш механизмига асосланган ахборот ишончилигини оптималлаштириш масаласи ечилган. Тадқиқотда, берилган мақсад функциясининг куйидаги ўзгарувчилари қийматлари мувофиқлаштирилади: u - танлама объектлари сони; t - ахборотни қайта ишлаш ва қидирув вақти; D - ўзгарувчи қиймати интервали; $\rho(t)$ - мақсад функциясига тўғри мурожат қилиш эҳтимоли; S - мақсад функциясига мурожаат сони; λ - мурожаат интенсивлиги; $\rho(u)$ - мурожаатнинг вақт функцияси. 1- расмда $\rho(u)$ ўзгарувчисининг турли қийматида берилган $\lambda(u)$ эҳтимол функцияси графиклари намоён қилинган.



1-расм. Таҳлика мақсад функцияси графиклари намоёни.

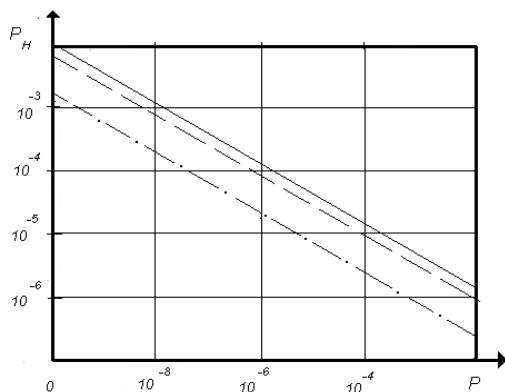
Қора чизик умумлашган алгоритмни қўллашда содир этиладиган йўқотиш эҳтимоли қийматини, кўк чизик анъанавий усулни, қизил чизик эса

алгоритмнинг оптималлаштириш механизмларини қўллаш вариантини акслантиради. Умумлашган алгоритмнинг хато тузатиш тўғри вариантга келгунча бўлган итерация сонини ўртача 5,6 маротобага қисқартиришга қодирлиги аниқланди. Жумладан, алгоритм матнда ҳарфлар алмашуви, қўшимчаси пайдо бўлиши, йўқотилиши, қўшни ҳарфлар алмашуви синфларига тегишли хатоларни бартараф қилади ва унинг самарадорлиги пакетли ва интерактив режимларда намоишланади.

Диссертациянинг «Матнда таркиблаштирилган информация ортиқчалигидан луғатсиз фойдаланувчи механизмлар бўйича ахборот ишончлилигини ошириш» номли иккинчи бобида матн n -грамм таркиблаштирилган моделлар, парсинг моделлаштириш, хэш - кодлаш механизмларига асосланиб тўламатли хужжатлар ахборот ишончлилигини оширишга йўналтирилган конструктив ёндашув ва алгоритмлар ишлаб чиқилган. Дарахтсимон қидирув модели, матн имловий хатоларини назорат, аниқлаш ва таҳрир қилиш механизмлари бўйича ҳамда қидириладиган сўзнинг сўз шакллари луғати рўйхатида мавжудлигини текширувчи алгоритмлар таклиф этилган. Матнни сатрнинг чап ёки ўнг тарафидан қидирувчи ҳамда матн сатрини ҳам чап, ҳам ўнг тарафдан навбат бўйича кузатувчи икки тарафли қидириш механизмларига асосланган модел ва алгоритмлар таклиф этилган. Натижада, алгоритмлар, сўзларни ажратиш, синфларга бўлишда тезкор ҳамда кам йўқотишга қодир бўлади. Имловий хатоларни назорат ва таҳрир қилиш механизмлари фреймворк «Сфинкс-4» ва «CUDA» параллел ҳисоблаш технологиялари муҳитига мослаштирилган ва ўзбек тилида берилган 5 томли оммавий адабиёт ва 16 томли сиёсий адабиётлар матнларидан ажратилган тўғри сўз корпуслари бўйича тестлаштирилган.

Бир каррали, n -грамм кўринишдаги хатолар содир бўлиш қонунияти ва шартли эҳтимолларидан фойдаланишга асосланган ахборот ишончлилигини оширувчи масала ечилган ҳамда алгоритмлари ишлаб чиқилган. Матн сатрида символлар пайдо бўлиши шартли эҳтимоли ҳисобланиши

$$p(\alpha_1 \dots \alpha_n) = \prod_{k=1}^n p(\alpha_k | \alpha_1 \dots \alpha_{k-1}) ; \log P(\alpha_1 \alpha_2 \alpha_3 \dots \alpha_n) = \sum_{k=1}^n \log P(\alpha_k | \alpha_1 \alpha_2 \alpha_3 \dots \alpha_n) \text{ энг қулай}$$



2-Расм Алгоритм самарадорлиги

бўлган қоида сифатида берилган. Ахборот ишончлилигини назорат қилиш алгоритми самарадорлиги диграммли содда модел кўринишидаги эҳтимоллар баҳоси бўйича тадқиқ қилинган

$$P(\alpha_i | \alpha_j) = \frac{F_{i|j}}{\sum_i F_{i|j}} = \frac{F_{i|j}}{F_j},$$

бунда F_i - символ ёки сўз частотаси; $F_{i|j}$ - i символдан кейинги j символлар шартли частотаси $P(\alpha_i | \alpha_j)$, яъни диграмм кўринишдаги шартли эҳтимоллар. Ушбу

таксимот модели $P(\alpha_i|\alpha_j) = \lambda \frac{F_i}{N} + (1-\lambda) \frac{F_{ij}}{F_i}$, N -символлар умумий сони; λ -

қиймати эмпирик равишда бериладиган параметр бўйича интерполяциялаштирилган. Алгоритм самарадорлиги P_H хатолар аниқланмаслиги эҳтимол мезони бўйича турли диграмм, учграмм, n -грамм хатолар шартли эҳтимоллари баҳолари бўйича таҳлил қилинган. 2-расмда ўрта эҳтимол хатонинг $P \approx 10^{-3} - 10^{-4}$ турли қийматларида функция P_H графиклари намоёниш қилинган. Ясси чизик диграмм, штрих чизик учграмм, штрих пунктирли чизик n -грамм эҳтимоллар формулалари асосида қурилган. Жумладан, бешграмм хатоликлар эҳтимолида алгоритм самарадорлиги икки тартибгача оширилиши аниқланган.

Қиёфаларни таниш, синфлаштириш, қидириш, n -грамм моделлардан фойдаланиш бўйича сўзларни эквивалентли классификаторларга ажратиш, парсинг кодлаштириш, сўз элементларини ҳамда сўзларни берилган кетма-кетликда эвристик ва параллел қидириш, ишончилиликни назорат қилиш функцияларини амалга оширувчи механизмлар асосида умумлаштирилган алгоритм ишлаб чиқилган. Умумлашган алгоритм сўз ишончилиги кўрсаткичини 5,6 маротобгача кўтаради, ахборотга ишлов бериш тезлигини эса кўп босқичли морфологик таҳлил моделига асосланган умумлашган алгоритмга нисбатан 2 маротобга оширади. Умумлашган алгоритм назорат қилинувчи сўзни таниш учун дарахтсимон моделга асосланган. Матн ишончилигини оширувчи дастурий мажмуа жорийлаштирилган ҳамда ўргатувчи ахборот векторининг турли узунлигида тестлаштирилган. Унинг қуйидаги хусусиятлари аниқланган: жуфт бўлмаган каррали хатоларни; монотонли хатоларни; симметрик хатоларни аниқлайди, ҳамда $d < M$ каррали монотонли хатоларни $d < M + 2$ каррали ассиметрик хатоларни тузатади.

Умумлашган алгоритм «Сфинкс-4» қобиғида, тўрт ядроли процессорда, матн имловий хатоларини назорат ва таҳрир қилувчи ДМ кўринишида яратилган ва жорийлаштирилган.

ДМ имконияти n -грамм парсинг моделлаштириш, кодлаш, қидириш механизмлари бўйича кенгайтирилган ҳамда матнли маълумотларга ишлов бериш билан боғлиқ турли хил масалалар ечимида синовдан ўтказилган. Умумлашган алгоритм хатоларни тузатиш керакли синфларда шакллантирилади, натижада итерация вақти 6 маротобгача қисқаради, шу жумладан, гипотезаларни текшириш, хотирага мурожаат, турли вариантда хатоларни тузатишга кетадиган вақтларни минималлаштиради.

Диссертациянинг «**Электрон хужжат матнлари таркибий қисмлари ишончилигини луғатсиз оширувчи механизмларни оптималлаштириш**» номли учинчи бобида чизик, текислик бўйича турли модуллар йиғиндиларини ҳосил қилиш, ортиқчаликка эга кодлашга асосланган ахборот ишончилигини луғатсиз назорат қилиш ёндашув, модел ҳамда алгоритмлари ишлаб чиқилган. Матнли маълумотларни кодлаш, сегментация ва назорат йиғиндиларини ҳосил қилиш услубияти таклиф этилган. Ишлаб чиқилган алгоритм хатоларни аниқлайди, ҳар бир гап

сўзларини тиниш белгилари бўйича ажратади, назорат қилинувчи сўзга морфологик хоссаларини бириктиради. Жумладан, алгоритм ҳар бир сўз фрагментига боғланиш турини белгилайди, сегментлар чегарасини мувофиқлаштиради ҳамда сегмент ичида кетма-кетликда синтактик информацияларни шакллантиради. Чизиқли, текислик бўйича модулли йиғиндилар ҳосил қилиш ва ортиқчаликка эга кодлаштириш механизмларига эга ахборот ишончилигини оширувчи алгоритмлар яратилган. Тадқиқотда, хатоларни аниқлаш ва тузатишда, ортиқча кодлаштиришга асосланган алгоритмларнинг катта самарадорлиги аниқланган. Жумладан, алгоритмлар назорат йиғиндиларидан фойдаланишда, матн сегментлари чегараларини мослаштиришда ижобий хоссаларга эга.

Алгоритмлар, айниқса, тиниш белгилари нотўғрилиги, ҳарфлар турли сурилишида, катта ва ёзув ҳарфлари нотўғри сурилганда, қўшни символ ва ҳарфлар ўрин алмашувида, икки каррали, кўп каррали ҳарфли, символли хатолар содир бўлганда ва йўқолганда ёки кўшимча ҳарф ва символлар пайдо бўлганда ҳамда форматли хатоларда самарадорлиги асосланди. Ишлаб чиқилган рақамли (ўнлик кодлаш) ва символли модуллар йиғиндиларини ташкил қилувчи механизмларга эга алгоритм, айниқса, n - каррали хатоларни аниқлашда катта қобилиятга эга. Бунинг учун қуйидаги шартлар бажарилиши талаб этилади. Хусусан, бир каррали хатоларни аниқлашда $\pm x \equiv r_a \pmod{q}$;

икки каррали хатоларни аниқлашда рақамлар йиғиндиси $\sum_{i=0}^n a_i$, $\pm x_2$,

$0 < x_2 \leq 18$; N сони ўзгарувчан катталиқ бўлганда; $x \cdot \omega \equiv r'_a \pmod{q}$ шарт

бажарилганда. Алгоритм $\sum_{i=0}^n a_i \omega_i$ амали имкониятини кенгайтириш бўйича

икки каррали хатоларни аниқлаш учун такомиллаштирилган. Бу ҳолда, алгоритмнинг ҳарф, рақам ва символлар ўрин алмашуви кўринишдаги хатоларни аниқлашда катта қобилиятга эгаллиги асосланган. Жумладан, алгоритм $10^i \equiv r'_{a_i} \pmod{q}$; $\pm x \equiv r'_{a_2} \pmod{q}$; $r'_{a_1} \cdot r'_{a_2} \equiv r'_a \pmod{q}$ q модул йиғиндилари бўйича такомиллаштирилган. Бу эса, яна $\pm x_i \cdot 10^{k_i} \pm x_{i-1} \cdot 10^{k_{i-1}} \pm \dots \pm x_1 \cdot 10^k = r'_a \pmod{q}$, $r'_a \neq 0$ шартларда алгоритмнинг k каррали хатоларни аниқлаш ва тузатишда самарадорлигини кўрсатади. Ахборот ишончилигини оширувчи алгоритмлар чизиқли ва текислик бўйича йиғиндилар ҳосил қилиш механизмлари асосида такомиллаштирилган. Бу алгоритмлар қуйидаги ҳолатларда хатоларни аниқлай олмайди:

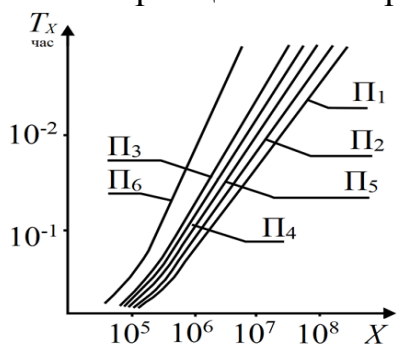
а) бир сатрда икки рақам ўрин алмашуви содир бўлганда ва бу ҳолат матн устунни рақамига мос келганда; б) сатрда бир рақам айрим катталиқка ўзгарганда ва бу сатрларда бошқа рақам қиймати ушбу катталиқда камайганда.

Ахборот ишончилиги разрядли, рақамли ва текислик бўйича йиғиндилар ҳосил қилувчи механизмларни қўллаганда юқори қийматларда эришиш мумкинлиги бўйича асосланган.

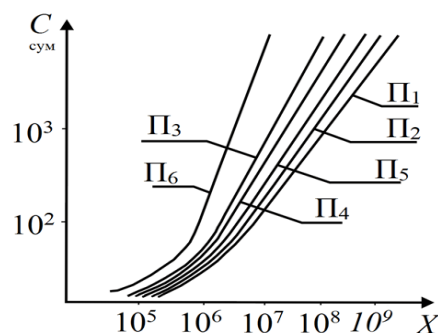
Алгоритм самарадорлиги 8, 9, 10, 11 ва ихтиёрий q модулли рақам, сон бўйича йиғиндилар ҳосил қилиш механизмларини қўллашда тадқиқ қилинган. Алгоритмнинг матнда турли хатоларни 92% гача аниқлаш ҳамда бир, икки каррали, ўхшаш ва транспозитив хатоларни тузатиш имкониятига эгаллиги исбот қилинган. Чизикли йиғиндилар ҳосил қилиш механизмида хатони аниқламаслик эҳтимоллар йиғиндиси шу ифода бўйича аниқланади.

$$P_H = \{67[mnp^2(1-p)^{m+n-2} + C_m^2 C_n^2 p^4(1-p)^{m+n-4} + C_m^k C_n^3 p^{3k}(1-p)^{m+n-k-3}] + 5,76 \cdot 10^{-6}\}.$$

3.1 ва 3.2 расмларда мураккаблиги ва ҳаражат нархи мезонлари бўйича ахборотни назорат қилиш самарадорлик графиклари келтирилган.



3.1. расм. Мураккаблик мезони бўйича алгоритм самарадорлиги



3.2. расм. Ҳаражат нархи мезони бўйича алгоритм самарадорлиги

Алгоритмларнинг чизикли, текислик бўйича йиғиндилар ҳосил қилувчи механизмларида ахборотни назорат қилиш мураккаблик ва ҳаражат нархи коэффициентлари қийматини 2,3 марта, ҳамда ишончилиликни иккидан уч тартибгача ошириш мумкинлиги аниқланган.

Диссертациянинг «Электрон хужжат алмашиш тизимида ахборот ишончилигини оширувчи дастурий мажмуани яратиш ва жорий қилиш» номли тўртинчи бобида гиперсемантик қидирув тармоғи, хужжат матнларида сўзларни таниш, синфлаштириш амалини бажарувчи воситаларга эга ахборот ишончилигини ошириш дастурий мажмуаси функционал модуллари ва дастурий таъминоти яратилган ҳамда қўлланилган. ЭХТ тўламанли хужжатлар ахбороти ишончилигини ошириш масалалари луғатсиз бажарилади, ДМ бир, икки, t - каррали хатоларни ортикчалikka эга кодлаш, чизикли, текислик ва модулар бўйича йиғиндилар ҳосил қилиш, фрейм, калитли жумла, термин, МБ ва ББ дан фойдаланувчи механизмлари асосида аниқлайди ва тузатади. Ўзбек тили матнлари орфографиясини назорат ва таҳрир қилиш учун кўп босқичли морфологик ва n -грамм таҳлил алгоритмлари, хатоларни аниқлаш ва тузатиш имконини берувчи услуб, процедуралар синфлари, кутубхонаси ишлаб чиқилган ва жорийлаштирилган. Кўп босқичли морфологик ва n -грамм таҳлил моделларига асосланган ДМ ахборот ишончилиги, яхлитлиги ва ҳимоясини таъминлашга қаратилган Web-йўналиш вариантыда қўлланилган анъанавий усуллар имконияти, чегараларини сезиларли даражада кенгайтиради. Дастурий таъминот объектга йўналтирилган дастурлаш технологияси

асосида яратилган, тадбикий ва дастурий модуллари синтезлаштирилган. Натижада, ташқи таъсирларга ДМ турғунлиги сезиларли даражада ошиши ҳамда априор маълумотлар танқислиги, ностационарлик, параметрлар ноаниқлиги мавжуд бўлган шароитларда назорат қилинувчи ахборот сифати юқори даражада таъминланиши асосланган. ДМ таркиби сўз шакли луғатини шакллантириш, *n*-грамм частотали луғат, иш юритиш термин (ибора) жамламаларини йиғиш ва улардан фойдаланиш, матнли ахборотларни уч тилга транслитерация қилиш, матн орфографиясини назорат ва таҳрир қилиш 12 функционал модуллардан иборат. Бундан ташқари, мажмуанинг функционал модулларидан мустақил фойдаланиш мумкин. Интегратор ва мослаш, маълумотларни умумлаштириш, импорт ва экспорт қилиш; матн семантикаси синтезатори ва анализатори; генерациялаш, сошлаш ва параметрларни мослаш; маълумотларни юклаш воситаси, интерфейс; калит сўз ва луғат кутубхонаси; маълумот конвертори ва ўгирувчиси; ТФХ тақсимлаш ДМ сервис модулларини ташкил қилади.

Дастур кутубхонаси қуйидагилардан иборат: меню, ЭҲга ишлов бериш графиги, матнни генерация, таҳлил, синтез, транслитерация қилиш ва бериш. Ўзбек тили матнли орфографияси назорат ва таҳрири фреймворк «Сфинкс-4» муҳитида амалга ошади ва бу ДМ семантик гипертармоқда қидириш, МБ, ББ, хэш – функциялар сонини, массив ўлчамини мувофиқлаштириш, *n*-грамм частотали луғатни шакллантириш, сўзни башорат қилиш, парсингни ҳисоблаш, сўз шакллари конструктори, эквивалентли синфлаш, таниш механизмларидан иборат. ДМ такомиллаштирилган варианты *n*-грамм таҳлил ва «CUDA» параллел ҳисоблаш технологиялари бўйича амалга оширилган.

Механизм ва технологияларни қўллаш, кўп босқичли морфологик таҳлил моделидаги вариант натижасига нисбатан хато сўз содир бўлишини аниқламаслик эҳтимолини уч тартибгагача камайтириши, мажмуа тезлигини эса икки маротабага ошириш имкониятига эгаллиги исбот қилинган. ДМ алгоритмлари самарадорлиги ахборотни назорат қилиш мураккаблик ва ҳаражат нархи мезонлари, параметрларни оптималлаштириш шохча, чегара ва стохастик моделлаштириш усуллари, чизикли дастурлашни қўллашдан олинган натижалар бўйича тадқиқ қилинган. 1-жадвалда тадқиқот предмет соҳасида қўллаш учун тавсия этилган ахборотни назорат қилувчи алгоритмлар самарадорлиги ишончлилиги бўйича ютуқ, мураккаблик ва нарх ҳаражати коэффициентлари асосида кўрсатилган.

Бажарилган қиёсий таҳлил яратилган алгоритмларнинг маълумотга пакетли ишлов бериш ва визуал назорат қилиш анъанавий тизими кўрсаткичларига нисбатан ахборот ишончлилиги икки-уч тартибга оширилиши; мураккаблик ва қиймат коэффициентларини беш-олти маротабага камайтирилиши ҳамда матндаги имловий хатоларнинг 92% гача бартараф қилиниши асосланган. Сирдарё вилояти ва Гулистон шаҳри турли корхоналари, маъмурий органлари фаолиятида ДМ ни қўллашда олинадиган иқтисодий самарадорликни ҳисоблаш услубияти яратилган. Ишлаб чиқилган, визуаллаштириш, трансформациялаш, таҳлил қилиш, компьютерда ишлов бериш алгоритмлари аниқ предмет соҳага қўлланишлиги учун тавсия этилган.

Ахборот ишончилигини ошириш алгоритмлари самарадорлигининг қиёсий таҳлили

№	Алгоритм	Самарадорлик мезонлари қийматлари								Алгоритм функциялари
		Хато аниқланмаслик эҳтимоли - P_H		Ютук коэффициенти K_{PH}		Мураккабликни камайтириш коэффициенти - K_T		Ҳаражат қийматини камайтириш коэффициенти - K_C		
1	2	3	4	5	6	7	8	9	10	11
	Анъанавий технология	$\approx 10^{-3}$	$4 \cdot 10^{-3}$	1	4	1	1	1	1	Визуал назорат
Ахборот ишончилигини лугатсиз оширувчи алгоритмлар										
	Чизиқли йиғиндилар	10^{-5}	$2 \cdot 10^{-5}$	100	200	1,8	1,7	2	2	Автоматлашტიрилган назорат
	Модулар бўйича йиғиндилар	$10^{-5}-10^{-6}$	$5 \cdot 10^{-6}$	1000	5000	2	2	2	2	- // -
	Текислик бўйича йиғиндилар	$10^{-5}-10^{-6}$	$1,5 \cdot 10^{-6}$	1000	1500	2	2	2	2	- // -
	Ортиқчаликка эга кодлаш	$10^{-5}-10^{-6}$	$4 \cdot 10^{-6}$	1000	4000	2	2	3	3	- // -
Ахборот ишончилигини лугат асосида оширувчи алгоритмлар										
	Кўп босқичли морфотаҳлил	$10^{-5}-10^{-6}$	$4 \cdot 10^{-6}$	1000	4000	2	2	3	3	- // -
	n-грамм модель	10^{-6}	$2 \cdot 10^{-6}$	1000	2000	3	3	3	3	- // -

ХУЛОСА

«Вилоят бошқарувида «электрон ҳукумат» тизими ҳужжатларига ишлов бериш ишончилигини ошириш усуллари ва алгоритмлари» мавзусидаги диссертация ишида олинган тадқиқот натижалари қуйидаги хулосаларни тақдим этиш имконини беради:

1. Сунъий, табиий, таркибий – технологик информация ортиқчаликлардан фойдаланишга асосланган, ҳужжатлар ахбороти ишончилигини ошириш учун ишлаб чиқилган, модел, алгоритм ва оригинал функционал модуллар дастурий мажмуа сифатида қўлланилишда замонавий ахборот узатиш протоколлари таркибида назорат қилувчи воситалар имкониятини сезиларли даражада оширади.

2. Сўз шакллари умумий луғати ва n -грамм таснифлар бўйича шакллантирилган рационал ўлчамдаги частотали луғатлардан фойдаланишга ҳамда кўп босқичли морфологик ва n -грамм таркибли таҳлил моделларига асосланган ўзбек тили матнлари орфографиясини назорат ва таҳрир қилувчи алгоритмлар таклиф этилган.

3. Тўла матнли ҳужжат ахбороти ишончилигини оширувчи модел ва алгоритмлар қуйидаги механизмларни қўллаш асосида яратилган: сўз коди, элементлари, информация яқинлигини аниқловчи; разрядли, рақамли чизиқли йиғиндилар ҳосил қилувчи; турли каррали хатоларни аниқлаш ва тузатувчи; модулли йиғиндилар ҳосил қилувчи; ўзбек тили матни кўп босқичли морфотаҳлил ва n -грамм таркибий модулларини қўлловчи. Ушбу механизмларга асосланган алгоритмлар харф ва сўзларни ахтаради, хато элементларни аниқлайди ҳамда тузатади, синфлаштиради, частотали таснифларини фойдаланиш учун таркиблаштиради.

4. Ҳужжат элементлари информация яқинлиги, статистик, корреляция, мантиқий, семантик боғланишлари асосида матн ахбороти ишончилигини оширувчи алгоритмлар таклиф этилган ва жорийлаштирилган.

5. Таклиф этилган моделлар асосида ишлаб чиқилган алгоритмлар чизиқ, текислик бўйича, рақам, символлар модуллари йиғиндиларини ҳосил қилиш ва матнни ортиқчаликка эга кодловчи механизмларни қўллайди, хатоларни луғатсиз равишда аниқлайди ҳамда тузатади. Шу жумладан, нотўғри тиниш белгилари, турли алфавитли ҳарфлар сурилиши, катта ва ёзув ҳарфлари, символлар ўрин алмашуви кўринишидаги бир каррали хатоларни тезкор равишда бартараф этади.

6. Алгоритмлар параметрлари оптималлаштирилди ҳамда улар самарадорлиги хато аниқланмаслик эҳтимоли, назорат мураккаблиги ва ҳаражат нархи мезонлари бўйича тадқиқ қилинди, натижада, ахборот ишончилигини визуал назорат қилиш технологиясига нисбатан ишончилик икки тартибдан учгача, мураккаблик ва ҳаражат нархи коэффициентлари қийматлари 5, 6 маротоба камаяди, 92 % гача матн орфографик хатолари аниқланади, шу жумладан, барча кўринишдаги икки каррали, ўхшаш ва транспозитив хатолар тузатилади.

**НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.01 ПО ПРИСУЖДЕНИЮ
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

ГУЛИСТАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ТИШЛИКОВ СУЛТОНЖОН АБДУРАИМОВИЧ

**МЕТОДЫ И АЛГОРИТМЫ ПОВЫШЕНИЯ ДОСТОВЕРНОСТИ
ОБРАБОТКИ ДОКУМЕНТОВ СИСТЕМ «ЭЛЕКТРОННОГО
ПРАВИТЕЛЬСТВА» ОБЛАСТНОГО УПРАВЛЕНИЯ**

05.01.02 - Системный анализ, управление и обработка информации

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ**

Ташкент–2021

Тема диссертации доктора философии (PhD) по техническим наукам зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за В2021.4.PhD/T1296

Диссертация выполнена в Гулистанском государственном университете.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице научного совета (www.tuit.uz) и на Информационно-образовательном портале «ZiyoNet».

Научный руководитель: **Жуманов Исраил Ибрагимович**
доктор технических наук, профессор

Официальные оппоненты: **Бабомуратов Озод Жураевич**
доктор технических наук, старший научный сотрудник
Сейтов Айбек Жумабаевич
доктор философии по техническим наукам (PhD),
старший научный сотрудник

Ведущая организация: **Ташкентский государственный транспортный университет**

Защита диссертации состоится «14» сентября 2022 г. в 16⁰⁰ часов на заседании Научного совета DSc.13.30.12.2019. Т.07.01 при Ташкентском университете информационных технологий. (Адрес: 100202, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-43; факс: (99871) 238-65-52; e-mail: tuit@tuit.uz).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского университета информационных технологий (регистрационный номер № 233). (Адрес: 100202, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-65-44).

Автореферат диссертации разослан «29» сентября 2021 года.
(протокол рассылки № 44 от «20» сентября 2021 г.).



Р.Х.Хамдамов

Председатель научного совета по присуждению учёных степеней, доктор технических наук, профессор

Ф.М.Нуралиев

Ученый секретарь научного совета по присуждению учёных степеней, доктор технических наук, доцент

А.В.Кобулов

Председатель научного семинара при научном совете по присуждению учёных степеней, доктор технических наук., профессор

ВВЕДЕНИЕ (аннотация диссертации доктора философии PhD)

Актуальность и востребованность темы диссертации. В мире особое внимание уделяется глубокому изучению вопросов разработки методов и алгоритмов для комплексной автоматизации и цифровизации систем управления производственными процессами и развитию систем электронного документооборота (СЭД), которые осуществляются в направлении интеграции технологий обработки информации, разветвленных сетей управления через WEB, автоматизации офисных процессов. В настоящее время проводятся научные исследования по оптимизации обработки документированной информации, отраженные в большом числе публикаций международных компаний Gartner, International Data Corporation, Forrester, ЮНИПРАВЭКС, Интерфейс ИТ, Интеллект Телеком и в других с указанием путей развития автоматизированных систем управления производственно - технологическими комплексами, компьютерных сетей, информационных систем, которые органически включаются в СЭД административных органов области для принятия управленческих решений.

В мире ведутся интенсивные исследования по повышению достоверности информации и достигнуты определенные успехи учеными зарубежных стран США, Великобритании, России, Китая, Индии, Германии, Австрии, Швеции, Норвегии, Финляндии, Турции, Израиля, Малайзии, Канады, Кореи, Бразилии и др. Достоинством их результатов является то, что они дают возможность формирования научно - методических основ, методов, моделей а также инструментов для обеспечения достоверности, защищенности, целостности текстов электронных документов (ЭД) с построением механизмов обнаружения и исправления ошибок.

Правительство Республики Узбекистан предприняло меры по широкому развитию науки, образования, цифровой экономики, внедрения инновационных идей и технологий, новых методов обработки информации в различных предметных областях. Выделяется актуальность мер, направленных присоединению к ним также технологий повышения достоверности информации на всех этапах ввода, передачи, хранения и обработки документов, решения которых остро востребованы и применяются в качестве инструментов снижения влияний человеческого фактора. Особенностью проблемы является то, что в Республике вся деловая документация ведется на узбекском языке и не все граждане обладают умением правописания текстов и в результате чего появляются орфографические ошибки. Наряду с этим, низкая достоверность информации обуславливается ошибками человека - оператора, погрешностью сканирования и распознавания, помехами в системе передачи информации, неисправностями средств подготовки и обработки информации, а также орфографическими ошибками пользователей системы. Это исследование, направленное для обеспечения достоверности документированной информации должны способствовать на повышении мобильности,

производительности, устойчивости СЭД при условии априорной недостаточности, нестационарности, параметрической неопределенности. А также подходы к ним должны быть направлены на построении и использовании эффективных механизмов обнаружения и исправления ошибок различной кратности текстов при наименьших материальных и временных затратах, а также учитывать параметрическую неопределенность процесса передачи и обработки ЭД.

Данное диссертационное исследование направлено в определенной степени на выполнение задач, предусмотренных Указом Президента Республики Узбекистан от 7 февраля 2017 года №УП-4947 «О стратегии действий по дальнейшему развитию Республики Узбекистан», от 29 августа 2017 года №ПП-3245 «О мерах по дальнейшему совершенствованию системы управления проектами в сфере информационно-коммуникационных технологий», от 19 февраля 2018 года №УП-5349 «О мерах по дальнейшему совершенствованию информационно-коммуникационных отраслей», от 6 октября 2020 года № ПП-4851 «О мерах по дальнейшему совершенствованию системы образования в области информационных технологий, развитию и интеграции научных исследований с IT-индустрией» другими нормативно-правовыми документами, принятыми в данной сфере.

Соответствие исследования приоритетным направлениям развития науки и технологии республики. Данное исследование выполнено в соответствии с приоритетным направлением развития науки и технологий Республики Узбекистан IV. «Развитие информатизации и информационно-коммуникационных технологий».

Степень изученности проблемы. Аналитическое изучение научно-технической литературы за последние 10-15 лет показало, что проблемам повышения достоверности текстовой информации в СЭД, основывается на исследования и разработки, направленные на создании систем морфологического анализа посвящены работы зарубежных ученых Г.Г.Белоногова, С.А.Старостина, М.Г. Мальковского, Кемаля Офлазера, Киммо Коскенни. Большой вклад в развитии теории и практики обработки документированной информации в системах управления внесли ученые Ch.W.Bachman, W.W.Peterson, D.Ritchi, E.F.Codd, C.J.Date, Alan Turing, E.N.Gilbert, R.C.Bose, Danish Jamil, Peter Pin-Shan Chen, Чжо Зо Е, В.Б.Кудрявцев, Ю.В. Семенов, В.Р.Сабанин, В.В. Трофимов, М.В. Ларин, Г.Ю Максимович, Ю.В. Гуляев и др.

Значительный вклад в развитии теории оптимальной обработки информации внесли ученые в нашей республике Д.А.Абдуллаев, Ф.Б.Абуталиев, М.М.Камилов, Т.Ф.Бекмуратов, Х.З.Игамбердиев, Ш.Х.Фазилов, И.И.Жуманов, Р.Х.Хамдамов, М.А.Рахматуллаев, А.А. Саидов А.Р. Ахатов и др. Определено, что технологии, гарантирующие достоверность информации при условии априорной ограниченности, неопределенности и использования механизмов извлечения свойств данных и

моделей описания вероятностных процессов недостаточно изучены, развитию и применению их пока находятся на начальной стадии.

Большую значимость имеет исследование и разработка словарного и без словарного методов, моделей и алгоритмов повышения достоверности информации в СЭД на основе применения механизмов многоуровневого морфологического и структурированного n -грамм анализов текстовых представлений посимвольного, цифрового, плоскостного суммирования, избыточного кодирования по модулям 8, 9, 10, 11 и q , обнаружения и коррекция орфографических ошибок с использованием искусственной, естественной, семантической, структурно – технологической информационной избыточностей с функцией автоматизированного устранения одно-, дву-, три-, n - кратных ошибок.

Связь темы диссертации с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация. Диссертация выполнена в рамках планов НИР СамГУ, ГулГУ, соответствующих приоритетному направлению НТП-17: «Разработка современных информационных систем, интеллектуальных средств управления и обучения, научно-технических баз данных и программных продуктов, обеспечивающих широкое развитие и внедрение информационных и телекоммуникационных технологий» (2009-2011) и ППИ-5: «Разработка информационных технологий, телекоммуникационных сетей, аппаратно-программных средств, методов и систем интеллектуального управления и обучения, направленных на повышение уровня информатизации общества» (2012-2014).

Цель исследования. Разработка и реализация моделей, алгоритмов, программных комплексов повышения достоверности документированной информации в системах «электронного правительства» административных органов областного управления.

Задачи исследования:

разработка методических основ, моделей, алгоритмов, используемых с механизмами использования словарного и без словарного контроля достоверности информации, многоуровневого морфологического и структурированного n -грамм методов анализа текстов на узбекском языке;

разработка и реализация моделей и алгоритмов обнаружения и исправления ошибок в текстах с механизмами линейного, плоскостного, модульного суммирования и избыточного кодирования;

оптимизация параметров алгоритмов повышения достоверности информации, исследование эффективности их по критериям вероятности необнаруженных ошибок, трудоемкости и стоимости контроля информации;

разработка функциональных модулей, общего программного обеспечения, программного комплекса повышения достоверности информации и реализация его в СЭД «электронного правительства» административных органов областного управления.

Объект исследования. СЭД, деловые документы, используемые в деятельности предприятий, организаций, учреждений и административных органов.

Предмет исследования. Модели, алгоритмы, инструменты обнаружения и исправления ошибок, механизмы линейного, модульного, плоскостного суммирования, избыточного кодирования, многоуровневого морфологического и n -грамм анализов.

Методы исследования теории передачи и обработки информации, алгоритмизации, математического программирования, использования алгоритмических языков высокого уровня, поиска и распознавания.

Научная новизна исследования заключается в следующем:

предложены конструктивные подходы, модели и алгоритмы повышения достоверности информации, основанные на использовании искусственной, естественной, семантической и структурно–технологической информационной избыточности;

разработаны механизмы многоуровневого морфологического и структурированного n -грамм анализов, линейного, плоскостного, модульного суммирования и избыточного кодирования текстов;

совершенствованы алгоритмы и разработаны механизмы обнаружения и исправления ошибок различной кратности на основе поиска с перебором всех вариантов, эвристического поиска с отжигом, запретами, со стохастическим моделированием на основе усеченной цепи Маркова;

оптимизированы параметры и исследована эффективность алгоритмов повышения достоверности информации по критериям достоверности, трудоемкости и стоимости контроля информации.

Практические результаты исследования состоит в следующем:

разработаны функциональные модули общего программного обеспечения, программный комплекс повышения достоверности информации;

реализована методика оптимизации параметров алгоритмов, механизмы формирования и использования частотных словарей n -грамм, обнаружения и коррекции ошибок различной кратности в текстах;

получены инструменты контроля орфографии текстов в соответствии требований ППП «Сфинкс-4» и параллельных вычислений «CUDA».

Достоверность результатов исследования. Достоверность научных результатов исследования подтверждается детальным математическим исследованием предложенных моделей, строгими математическими преобразованиями, сравнительным анализом, применением полученных формул и выкладок с реальными и экспериментальными данными на основе общепринятых критериев, компьютерным моделированием алгоритмов и реализацией разработанных программных комплексов.

Научная и практическая значимость результатов исследования заключается в следующем:

научная значимость результатов исследования заключается в том, что предложены модели, принципы и механизмы повышения достоверности информации со словарем словоформ с использованием различной природы информационной избыточности, многоуровневого морфологического и *n*-грамм структурированного анализа грамматики узбекского языка; без применения словаря, заменяемых механизмами линейного, плоскостного, модульного суммирования, избыточного кодирования текстов; оптимизации алгоритмов повышения достоверности информации по критериям вероятности необнаруженных ошибок, коэффициентам трудоемкости и стоимости контроля информации;

практическая значимость результатов исследования заключается в следующем. Предложена схема интеграции оригинальных функциональных модулей и разработан программный комплекс (ПК) повышения достоверности информации, основанные на механизмы применения словаря словоформ и без словарные механизмы обнаружения и исправления ошибок *n*-кратности, смешенного алфавита, пропуска, замены букв, а также другие орфографические ошибки текстов на узбекском языке.

Внедрение результатов исследования. На основе реализации результатов исследования, реализованных алгоритмов и программных средств оптимизации повышения достоверности передачи документов в системе «Электронного правительства» региона выполнены следующие экспериментальные работы:

программное обеспечение, инструменты, рекомендации, разработанные механизмы, повышающими достоверность обработки информации было внедрено в деятельности «УЗБЕКТЕЛЕКОМ» Сырдарьинский филиал «СырдарьяТелеком» (Справка Министерства по развитию информационных технологий и коммуникаций №33-8/136 от 8 января 2021 г.). В результате доказано, что разработанные модели, алгоритмы, программные средства, по сравнению механизма визуального контроля, значение показателя достоверности информации повышает на 2-3 порядков, по коэффициентам трудоемкости и стоимости контроля и обработки информации показывает на 5-6 раза уменьшения их значений; разработанные модели и реализованные программные средства внедрены в деятельности Главного управления занятости Сырдарьинской области (Справка Министерства по развитию информационных технологий и коммуникаций №33-8/136 от 8 января 2021 г.). Доказано, что значение показателя достоверности информации повышается на 2-3 порядка, а также они способствуют сокращению материальных расходов на 15% при исполнении организационно-распорядительных документов; разработанные и реализованные программные средства внедрены в деятельности «АПК Оқ олтин» УК (Справка Министерства по развитию информационных технологий и коммуникаций №33-8/136 от 8 января 2021 г.). Установлено, что программный комплекс позволяет обнаруживать до 92% орфографических ошибок и исправлять различных двукратных, смежных, транс позитивных

ошибок. Благодаря применению программного комплекса значения коэффициентов трудоемкости и стоимости контроля и обработки информации уменьшаются в 5-6 раза.

Апробация результатов исследования. Результаты данного исследования обсуждены на 11 международных и 2 республиканских научно-практических конференциях.

Публикация результатов исследования. По теме диссертации опубликовано 32 научных работ, из них 10 журнальных статей, в том числе 1 в международных и 9 в республиканских журналах, рекомендованных Высшей аттестационной комиссией Республики Узбекистан для публикации основанных научных результатов докторских диссертаций, также получены 4 свидетельства о регистрации программных продуктов для ЭВМ.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы и приложений. Объем диссертации составляет 112 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обосновываются актуальность и востребованность темы диссертации, формулируются цель, задачи, объект, предмет исследования, научная новизна, практические результаты исследования, обоснована достоверность полученных результатов, раскрыта теоретическая и практическая значимость результатов, приведены перечень внедрений, сведения об опубликованных работах и структуре диссертации.

В первой главе диссертации «**Механизмы повышения достоверности информации со словарем словоформ в системах управления**» разработаны конструктивные подходы, принципы, методы, направленные обнаружению и исправлению n -кратных ошибок на основе использования механизмов искусственной, естественной, семантической, структурно - технологической информационной избыточностей, многоуровневого морфологического и структурированного n -грамм моделей анализа текстов, линейного, плоскостного, модульного суммирования а также избыточного кодирования.

В настоящее время интенсивно ведутся разработка и реализация технологий, программных сервисов, информационно-ресурсных сетей в СЭД административных органов для принятия управленческого решения. Большую значимость представляет вычислительные схемы, построенные на основе многоуровневого морфологического и n -грамм моделей анализа текстов на узбекском языке.

Исследована статистика ошибок типа замена, вставка, пропуск, перестановка букв и символов и др. Разработаны модифицированные модели и алгоритмы повышения достоверности информации полнотекстовых документов с механизмами представления дополнительной информации, лингвистических знаний о каждой словоформе, а также с применением

инструментов грамматического, синтаксического, морфологического анализа, обнаружения и исправления ошибок n -кратности. Предложены алгоритмы повышения достоверности информации модифицированные на основе включения механизма эвристического поиска со стохастическим моделированием, которые придают им способности при оперативном обнаружении и исправлении наиболее распространенных однократных, транспозиционных, двукратных, смежных ошибок и ошибок высокой кратности, а также ошибок типа пропуск, добавления букв и символов.

Разработаны механизмы, которые способствуют совершенствованию традиционных систем контроля орфографии «spell-checker» узбекского языка, оптимизируют размера общего словаря словоформ и частотных словарей, исправлению однократных, двукратных, комбинированных ошибок. Эффективность алгоритмов повышается путем априорного сужения пространства поиска словоформ, упорядочения вариантов поиска с перебором, применения эвристики с отжигом, запретом, механизма извлечения свойства информации в базы знаний (БЗ). Безсловарных алгоритмы ярким образом проявляют свои преимущества при использовании накопленной и представленной статистики искажений, частотных словарей n - грамм характеристик текстов, шаблонов ЭД, а также модели совершаемых искажений в информации.

Разработаны алгоритмы адаптивного контроля достоверности информации, которые обладают способностью в параллельном исправлении монограммных и комбинированных ошибок. Построен обобщенный алгоритм для контроля и коррекции орфографии текстов, который совмещает механизмы гиперсемантической сети поиска, представления набора базовых фреймов, шаблонов и ключевых слов, регулирования числа поисковых объектов, времени поиска, интервала значений переменных, числа правильного и неправильного обращений к целевой функции, интенсивности внешних воздействий. Тестирование проведено по набору из 2000 фреймов 12 документов. Установлено, что обобщенный алгоритм, чем технология визуального контроля орфографии текстов позволяет сократить общее время контроля достоверности информации почти 6 раза.

Оптимизирована достоверности информации с реализацией и применением методики поиска глобального и локальных экстремумов на основе градиентного метода. Исследование проведено по функции риска, в котором предусмотрено регулирование следующих переменных: u - числа объектов в наборе; t - время поиска и переработки информации; D - интервал значений переменных; $\rho(t)$ - вероятность правильного обращения к целевой функции; S - числа обращений к целевой функции; λ - интенсивность обращения; $\rho(u)$ - время обращений.

На рис. 1 показаны графики вероятностной функции $\lambda(u)$, которые построены при различных значениях переменной $\rho(u)$.

Линией черного цвета отражаются значения вероятности риска при применении обобщенного алгоритма, синий цвет традиционного, линия красного цвета поведение целевой функции при механизме оптимизации.

Определено, что обобщенный алгоритм до получения варианта исправления ошибок позволяет уменьшить числа итераций в среднем 5, 6 раза. При этом исправляются ошибки типа замена, вставка, пропуск одной буквы, перестановка соседних букв. Алгоритм демонстрирует свою эффективность как при пакетном, так интерактивном режимах СЭД.

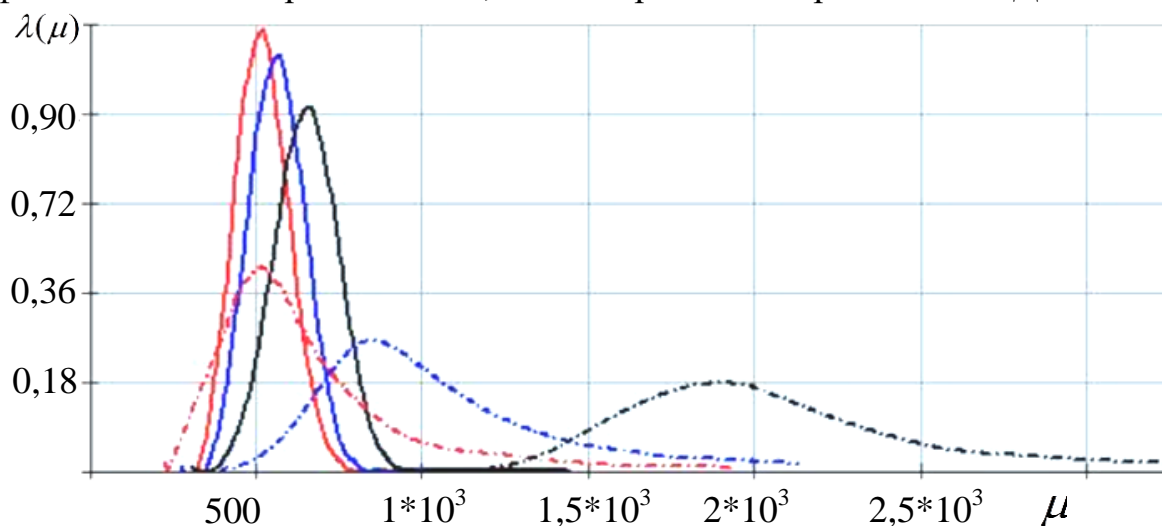


Рис. 1. Иллюстрация графиков поведения целевой функции риска.

Во второй главе диссертации «**Безсловарные методы повышения достоверности информации на основе механизмов использования структурной информационной избыточности в текстах**» разработаны конструктивные подходы, принципы и методы повышения достоверности информации полнотекстовых документов, направленные на применении механизмов структурированных n -грамм модели, парсинг моделирования, хэш – кодирования, проверки наличия слова в частотных словарях словоформ, формирования, представления, идентификации n -грамм, обнаружения и коррекции орфографии, поиска объекта по древовидной модели. Реализован подход, который направлен на использовании механизмов одностороннего просмотра строки текста слева или справа, двухстороннего поиска по каждой строке текста. Причем, строка текста поочередно прослеживается и слева и справа. Установлено, что такие механизмы алгоритмов позволяют быстрее и без заметных потерь выделять слова, разделяет его на классы. Алгоритмы контроля и коррекции орфографии текстов узбекского языка адаптированы в среде ППП «Сфинкс-4» и параллельных вычислений «CUDA». Тестирование обобщенного алгоритма проведено при достаточно обширном корпусе грамматически правильных слов пятитомного популярного литературного текста и корпуса слов из 16 томов политической литературы.

Разработаны алгоритмы повышения достоверности информации, которые основаны на использовании закономерностей появления однократных ошибок и условных вероятностей ошибок типа n -грамм. Условная вероятность символов в строке текста вычисляется по правилу

$$p(\alpha_1 \dots \alpha_n) = \prod_{k=1}^n p(\alpha_k / \alpha_1 \dots \alpha_{k-1}); \quad \log P(\alpha_1 \alpha_2 \alpha_3 \dots \alpha_n) = \sum_{k=1}^n \log P(\alpha_k | \alpha_1 \alpha_2 \alpha_3 \dots \alpha_n),$$

которое является удобной в вычислительном отношении. Эффективность алгоритма исследована на основе использования оценок вероятностных переходов в виде упрощенной модели определения диграмм. Условная вероятность

$$P(\alpha_i | \alpha_j) \text{ диграмм определена, как } P(\alpha_i | \alpha_j) = \frac{F_{i|j}}{\sum_i F_{i|j}} = \frac{F_{i|j}}{F_j},$$

или слова; $F_{i|j}$ - условные частоты, как число следования символа j за символом i . Распределение диграмм интерполировано, как

$$P(\alpha_i | \alpha_j) = \lambda \frac{F_i}{N} + (1 - \lambda) \frac{F_{i|j}}{F_j},$$

где N - общее число символов; λ - значения параметра определены эмпирическим способом. Эффективность алгоритмов проанализирована по критерию вероятности необнаруженных ошибок P_H при условных вероятностях диграмм, триграмм и n -грамм ошибок.

На рис. 2 показаны графики функции P_H , которые построены при значениях средней вероятности ошибок $P \approx 10^{-3} - 10^{-4}$. Сплошная линия при использовании вероятностей диграмм, штриховая триграмм, штрихпунктирная n -грамм. Эффективность алгоритма возрастает до двух порядков при вероятностях распределения пяти грамм. Разработан обобщенный алгоритм, направленный на использование инструментов распознавания, классификации, поиска и представления n -грамм, который эффективно выполняется при механизмах выделения слова в виде эквивалентных классификаторов, парсингового представления, кодирования, поиска слов, параллельного эвристического поиска. Определено, что обобщенный алгоритм улучшает показатель достоверности распознавания слово в 15, 16 раз, повышает скорость обработки информации по сравнению обобщенного алгоритма, основанного на механизм многоуровневой модели

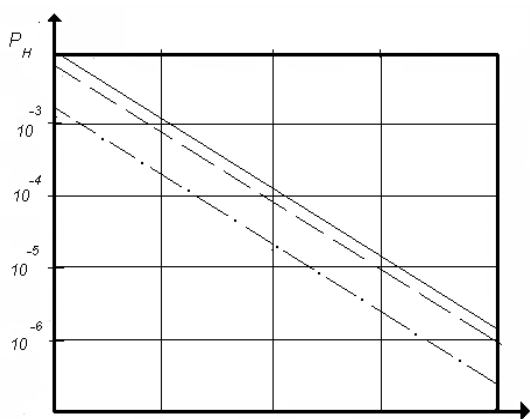


Рис.2. Эффективность обобщенного алгоритма

морфологического анализа 2 раза.

Обобщенный алгоритм реализован на основе парсингового дерева. Реализован программный комплекс для повышения достоверности текстов, который протестирован на различных длинах информационных векторов обучения. Выявлены следующие его особенности: обнаруживают ошибки с нечетными кратностями; монотонные ошибки; симметричные ошибки, а также

исправляют монотонные ошибки кратностью $d < M$ и асимметричные ошибки кратностью $d < M + 2$.

Программный комплекс реализован для решения разнообразных задач обработки текстовой информации в четырехядерном процессоре, который расширяет возможностей инструментов, входящих в оболочку «Сфинкс-4», адаптирует механизмов при применении модели n -грамм парсингового моделирования, кодирования, поиске. В результате осуществляет полный поиск, формирует требуемых классов, сокращает числа итерации, до 6 раз минимизирует времени проверки гипотез, обращений к памяти, нахождении всех вариантов исправленных ошибок.

Предложена методика перекодирования текстов, сегментации и образования контрольных сумм. Разработан алгоритм, направленный обнаружению ошибок в каждом предложении, применению механизмов разбиение слова на основе знаков препинаний, использованию свойства морфологической информации, соответствующей контрольному слово.

В третьей главе диссертации **«Оптимизация механизмов на основе без словарных методов повышения достоверности структурных составляющих текстов электронных документов»** разработаны основные подходы, модели и алгоритмы без словарного контроля достоверности информации с механизмами линейного, плоскостного, модульного суммирования и избыточного кодирования. Алгоритм каждому фрагменту слово приписывает тип связей, регулирует границ сегментов, формирует синтаксические информации. Разработаны алгоритмы повышения достоверности информации с механизмами линейного, плоскостного, модульного суммирования и избыточного кодирования. Определено, что среди них алгоритмы с механизмами избыточного кодирования обладают наибольшей эффективностью при обнаружении и исправлении ошибок. Положительные свойства алгоритмов установлены также при механизмах ведения контрольных сумм и регулирования границ сегментов текста. Алгоритмы эффективно исправляют ошибки типа неправильного знака препинания, смещения различных букв алфавита, неправильного смещения заглавных и прописных букв, одиночных буквенных и символьных ошибок, двукратных ошибок типа перестановки соседних букв и символов, многократных буквенных и символьных ошибок, пропуска или добавления букв и символов. Построен алгоритм с механизмом цифрового (десятичного кодирования) и посимвольного модульного суммирования, преимуществом которого является способность при обнаружении ошибок n -кратности. Для чего, при обнаружении одиночной ошибки должно выполняться условие $\pm x \equiv r_a \pmod{q}$, двукратных ошибок сумма цифр $\sum_{i=0}^n a_i$ числа N должно изменяться на величину $\pm x_2$, $0 < x_2 \leq 18$ и условие $x \cdot \omega \equiv r'_a \pmod{q}$.

Алгоритм модифицирован для обнаружения двукратных ошибок на основе расширения процедуры $\sum_{i=0}^n a_i \omega_i$. Алгоритм обладает способностью при обнаружении ошибок типа «перестановка» букв, цифр и символов. Алгоритм с механизмом суммирования по модулю q задается, как

$$10^i \equiv r'_{a_i} \pmod{q}; \pm x \equiv r'_{a_2} \pmod{q}; r'_{a_1} \cdot r'_{a_2} \equiv r'_a \pmod{q}.$$

Алгоритм модифицирован для обнаружения и исправления ошибок кратности k при условии $\pm x_i \cdot 10^{ki} \pm x_{i-1} \cdot 10^{k(i-1)} \pm \dots \pm x_1 \cdot 10^k = r'_a \pmod{q}$, $r'_a \neq 0$. Алгоритмы модифицированы также на основе механизмов линейного и плоскостного суммирования. Эти алгоритмы не обнаруживает ошибки при ситуациях, когда:

а) происходит перестановка двух цифр в одной строке и соответствующим этим цифрам столбцах;

б) одна цифра в строках увеличивается на некоторую величину, а другая цифра в этих строках уменьшается на ту же величину.

Определено, что высокое значение достоверности информации достигается при использовании механизмов поразрядного, цифрового и плоскостного суммирования. Эффективность алгоритмов исследована также при механизмах цифрового, числового суммирования на базисе модулей 8, 9, 10, 11 и произвольного q . Установлено, что при этом алгоритмы обнаруживают до 92% всех видов текстовых ошибок и исправляют в текстах одно, двукратные, смежные и транспозиционные ошибки. Суммарная вероятность необнаруженных ошибок при механизме линейного суммирования определяется, как

$$P_H = \{67[mnp^2(1-p)^{m+n-2} + C_m^2 C_n^2 p^4(1-p)^{m+n-4} + C_m^k C_n^3 p^{3k}(1-p)^{m+n-k-3}] + 5,76 \cdot 10^{-6}\}.$$

На рис. 3.1 и 3.2 показаны графики эффективности исследованных алгоритмов по критериям трудоемкости и стоимости контроля информации.

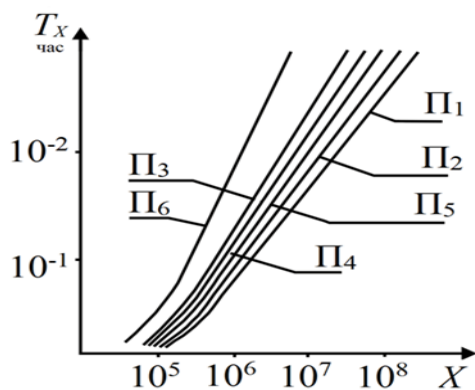


Рис. 3.2. Эффективность алгоритма при критерие трудоемкости

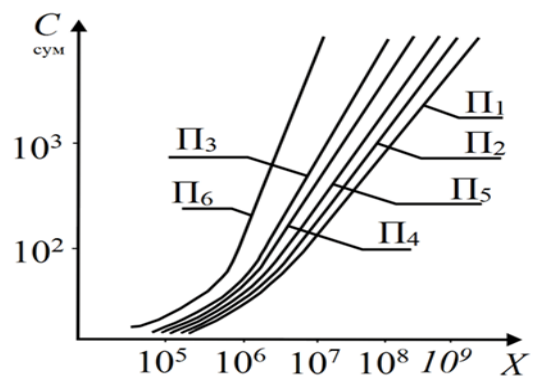


Рис. 3.2. Эффективность алгоритма при критерие стоимости

Определено, что алгоритмы при механизмах линейного и плоскостного суммирования значения коэффициентов трудоемкости и стоимости контроля информации снижает 2, 3 раза и повышают достоверность информации от двух до трех порядков.

В четвертой главе диссертации **«Разработка и реализация программного комплекса повышения достоверности информации в системе электронного документооборота»** разработаны и реализованы функциональные модули ПК повышения достоверности информации, программное обеспечение, которое включает модули гиперсемантической сети поиска, инструментов распознавания, классификации слов.

Реализованы функциональные модули повышения достоверности информации полнотекстовых документов без применения словаря словоформ. Протестированы также программные модули, которые способны обнаруживать и исправлять одно-, дву-, t -кратных ошибок на основе использования механизмов избыточного кодирования, линейного, плоскостного, модульного суммирования, представления базовых фреймов, ключевых фраз, терминов. Для контроля и коррекции орфографии текстов на узбекском языке реализованы механизмы многоуровневого морфологического и n -грамм анализов, методы, классы и библиотеки процедур, позволяющие повысить эффективность ПК при обнаружении и коррекции ошибок. Доказано, что ПК в Web-ориентированном варианте с механизмами многоуровневого морфологического и n -грамм анализов расширяют рамки, традиционных методов обеспечения достоверности, целостности и защищенности информации.

Программное обеспечение разработана на основе технологии объектно-ориентированного программирования, которое обобщает прикладные и функциональные модули. БД и БЗ придают ПК большой устойчивости к различным возмущениям, требуемой способности обеспечению высокого качества информации при условии априорной недостаточности, нестационарности, параметрической неопределенности.

ПК реализован в составе 12 функциональных модулей, выполняющие следующих функций: формирования и использования общего словаря словоформ, частотных n -грамм словарей, терминов делопроизводства, выполняющие транслитерации текстовой информации, контроля и коррекции орфографии текстов. Каждый функциональный модуль может использоваться автономно. Сервисными модулями ПК являются: интегратор и средства настроек, объединений, импорта, экспорта данных; семантический синтезатор и анализатор текста; средства генерации, настройки и адаптации параметров; интерфейс со средствами загрузки данных; библиотека ключевых слов и словарей; конвертор и преобразователь данных; распределитель ОРД.

Библиотека программ ПК включает: меню программ, графиков обработки ЭД, генерации, анализа, синтеза, транслитерации текстов и вывода. Параметры алгоритмов оптимизированы по методам ветвей, границ

и стохастического моделирования на основе линейного программирования. Адаптированный для контроля и коррекции орфографии в среде фреймворк «Сфинкс-4» и технологии параллельных вычислений «CUDA» ПК расширен на основе включения механизмов гиперсемантического поиска, БД, БЗ, определения количества хеш-функций, регулирование длины массива, формирования n -грамм частотных словарей, предсказателя слов, вычислителя парсингов, конструктора, формирования эквивалентных классификаторов, распознавателя словоформ. Доказано, что благодаря применению этих технологии и механизмов по сравнению варианта многоуровневого морфологического анализа вероятность необнаружения ошибочного слово уменьшается на три порядка и два раза повышается скорость.

Эффективность ПК исследована по критериям стоимости и трудоемкости контроля информации. В табл. 1 приведены результаты сравнительного анализа эффективности рекомендованных алгоритмов, которые отражаются по значениям коэффициентов выигрыша в достоверности, трудоемкости и стоимости контроля информации.

Определено, что алгоритмы по сравнению технологии визуального контроля, достоверность информации повышают до трех порядков; коэффициенты трудоемкости и стоимости контроля информации снижаются в 5-6 раза, а также до 92% устраняются орфографические ошибки в текстах документов.

Разработана методика расчета экономической эффективности внедрения ПК в СЭД предприятий и учреждений Сирдарьинской области. Описание возможностей программных средств, модулей и комплексов изложено в соответствующих аннотациях свидетельств Агенства интеллектуальной собственности РУз. В приложениях даны свидетельства о регистрации программ на ЭВМ.

Таблица 1

Сравнительный анализ эффективности алгоритмов ПК повышения достоверности информации

№	Названия алгоритмов	Значения критериев эффективности алгоритмов								Выполнение функции
		Вероятности необнаруженных ошибок - P_H		Коэффициент снижения P_H		Коэффициент снижения трудоемкости - K_T ,		Коэффициент снижения стоимости - K_C		
1	2	3	4	5	6	7	8	9	10	11
	Традиционная технология	$\approx 10^{-3}$	$4 \cdot 10^{-3}$	1	4	1	1	1	1	Визуальный контроль
Безсловарные алгоритмы повышения достоверности информации										
	Линейные суммирование	10^{-5}	$2 \cdot 10^{-5}$	100	200	1,8	1,7	2	2	Автоматический
	Модульные суммирования	$10^{-5}-10^{-6}$	$5 \cdot 10^{-6}$	1000	5000	2	2	2	2	- // -
	Плоскостное суммирование	$10^{-5}-10^{-6}$	$1,5 \cdot 10^{-6}$	1000	1500	2	2	2	2	- // -
	Избыточное кодирование	$10^{-5}-10^{-6}$	$4 \cdot 10^{-6}$	1000	4000	2	2	3	3	- // -
Словарные алгоритмы повышения достоверности информации ЭД										
	Многоуровневый морфоанализ	$10^{-5}-10^{-6}$	$4 \cdot 10^{-6}$	1000	4000	2	2	3	3	- // -
	n -граммная модель	10^{-6}	$2 \cdot 10^{-6}$	1000	2000	3	3	3	3	- // -

ЗАКЛЮЧЕНИЕ

Результаты диссертационной работы, полученные по теме «Методы и алгоритмы повышения достоверности обработки документов систем «электронного правительства» областного управления» позволяют заключить следующие выводы:

1. Разработанные модели, алгоритмы и оригинальные функциональные модули при повышении достоверности информации документов применяются на основе использования искусственной, естественной, структурно - технологической избыточностей в виде программных комплексов, которые позволяют существенно расширить возможностей инструментов контроля современных протоколов передачи данных.

2. Предложены алгоритмы контроля и коррекции орфографии текстов узбекского языка на основе многоуровневого морфологического анализа и n -грамм структуры с использованием рационального размера словаря узбекских словоформ и частотных словарей, полученных по частотным характеристикам n -грамм. Определена статистика характерных n -кратных искажений в текстах, использование механизма которых позволяет повысить эффективность обнаружения и коррекции ошибок различной кратности.

3. Разработаны модели, алгоритмы повышения достоверности информации полнотекстовых документов с использованием механизмов: определения информационной близости элементов текста, кодовых слов; обнаружения и исправления ошибок различной кратности на основе поразрядного, цифрового линейного суммирования; модульных суммирования, использования многоуровневой морфологической анализ и n -грамм структурированной моделей текстов узбекского языка. Алгоритмы с предложенными механизмами выполняют поиск и кластеризации, выявляют искаженных элементов (букв, слов) и осуществляют структуризации частотных характеристик.

4. Предложены и реализованы алгоритмы повышения достоверности текстов на основе использования информационной близости, статистических, корреляционных, логических, семантических связей элементов документов. Определено, что механизмы, при этом, способствуют выявлению схожих слов, уникальных свойств текстов, уникальных лексических единиц информации, использованию статистических связей элементов в наборах обучения, в базовых функциях инициализации всех текстов, логических связей элементов, БЗ, набора правил по контролю достоверности информации.

5. Разработаны модели и алгоритмы безсловарного обнаружения и исправления ошибок на основе механизмов линейного, плоскостного, цифрового, числового модульного суммирования и избыточного кодирования, которые позволяют оперативно устранять ошибки типов неправильного знака препинания, смещения букв различного алфавита,

заглавных и прописных букв, одиночных буквенных и символьных однократных ошибок, перестановки символов, замещений.

6. Оптимизированы параметры алгоритмов и исследована их эффективность по критериям вероятности необнаруженных ошибок, трудоемкости и стоимости контроля информации. Определено, что по сравнению с визуальным контролем достоверность информации повышается от двух до трех порядков, значения коэффициентов трудоемкости и стоимости контроля информации уменьшаются 5-6 раз, до 92% обнаруживаются орфографические ошибки текстов и исправляются все виды двукратных, смежных и транспозиционных ошибок.

**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES
DSc. 13.30.12.2019.T.07.01 AT TASHKENT UNIVERSITY OF
INFORMATION TECHNOLOGIES**

GULISTON STATE UNIVERSITY

TISHLIKOV SULTONJON ABDURAIMOVICH

**METHODS AND ALGORITHMS TO INCREASE THE RELIABILITY OF
DOCUMENT PROCESSING SYSTEMS OF "ELECTRONIC
GOVERNMENT" OF THE REGIONAL ADMINISTRATION**

05.01.02 – System analysis, management and information processing

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PhD)
ON TECHNICAL SCIENCES**

Tashkent–2021

The theme of dissertation of Doctor of Philosophy (PhD) on technical sciences was registered at the Supreme Attestation Commission at the Cabinet of Ministers of the Republic of Uzbekistan under number B2021.4.PhD/T1296

The dissertation has been prepared at the Guliston Staty University.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on the website of Scientific Council (www.tuit.uz) and on the website of «Ziyonet» Information and educational portal (www.ziyonet.uz).

Scientific adviser:

Jumanov Isroil Ibragimovich

Doctor of Technical Sciences, Professor

Official opponents:

Babomuratov Ozod Juraevich

Doctor of Technical Sciences,
Senior Research Fellow

Seytov Aybek Jumabaevich

Doctor of Philosophy (PhD) in Technical Sciences,
Senior Research Fellow

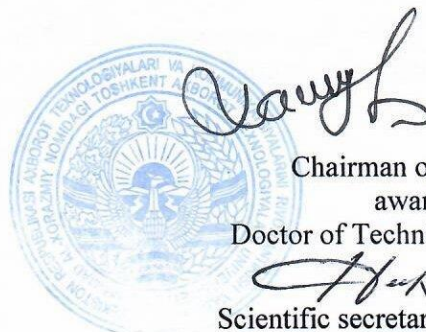
Leading organization:

Tashkent State Transport University

The defense will take place «14» January 2022 at 16⁰⁰ the meeting of Scientific Council № DSc. 13/30.12.2019.T.07.01 at Tashkent University of Information Technologies (Address: 100202, Tashkent city, Amir Temur street, 108. Tel.: (99871) 238-64-43, fax: (99871) 238-65-52, e-mail: tuit@tuit.uz).

The dissertation can be reviewed at the Information Resourse Centre of the Tashkent University of Information Technologies (is registered under № 138). (Address: 100202, Tashkent city, Amir Temur street. 108. Tel.: (99871) 238-64-43, Fax: (99871) 238-65-52).

Abstract of the dissertation sent out on «29» December 2021 y.
(mailing report № 44 on «20» December 2021 y.).



R. Kh. Khamdamov

Chairman of the Scientific Council
awarding scientific degrees,
Doctor of Technical Sciences, Professor

F. M. Nuraliev

Scientific secretary of Scientific Council
awarding scientific degrees,
Doctor of Technical Sciences, Docent

A. V. Kabulov

Chairman of the academic seminar under the
Scientific Council awarding scientific degrees,
Doctor of Technical Sciences, Professor

INTRODUCTION (abstract of PhD dissertation)

The aim of the research work. Development and implementation of models, algorithms, software systems for increasing the reliability of documented information in the systems of "electronic government" of regional administrative bodies.

The object of the research work. Texts, documents used in the activities of enterprises, institutions and administrative bodies.

The scientific novelty of the research work is as follows:

proposed constructive approaches, models and algorithms for increasing the reliability of information based on the use of artificial, natural, semantic and structural-technological information redundancy;

mechanisms of multilevel morphological and structured n-gram analyzes, linear, planar, modular summation and redundant coding of texts have been developed;

algorithms have been improved and mechanisms have been developed for detecting and correcting errors of various multiplicity based on search with enumeration of all options, heuristic search with annealing, prohibitions, with stochastic modeling based on a truncated Markov chain;

the parameters were optimized and the efficiency of algorithms for increasing the reliability of information was investigated according to the criteria of reliability, labor intensity and cost of information control.

Implementation of the research results. Based on the results of the research implementation, the implemented algorithms and software tools for optimizing the increase in the reliability of the transmission of documents in the "Electronic Government" system, the following experimental works were performed:

software, tools, recommendations, developed mechanisms that increase the reliability of information processing were introduced in the activities of UZBEKTELECOM Sirdarya branch of SirdaryaTelecom (Reference of the Ministry for the Development of Information Technologies and Communications No. 33-8 / 136 dated January 8, 2021). As a result, it has been proved that the developed models, algorithms, software, by comparing the mechanism of visual control, the value of the information reliability indicator increases by 2-3 orders of magnitude, in terms of labor intensity and cost of control and information processing, it shows a 5-6 times decrease in their values;

the developed models and implemented software tools have been introduced in the activities of the Main Employment Department of the Syrdarya region (Reference from the Ministry for the Development of Information Technologies and Communications No. 33-8 / 136 dated January 8, 2021). It has been proved that the value of the information reliability indicator increases by 2-3 orders of magnitude, and they also help to reduce material costs by 15% in the execution of organizational and administrative documents

the developed and implemented software tools have been introduced in the activities of "APK O'Oltin" UK (Reference from the Ministry for the Development

of Information Technologies and Communications No. 33-8 / 136 dated January 8, 2021). It has been found that the software package can detect up to 92% of spelling errors and correct various double, adjacent, trans positive errors. Thanks to the use of the software complex, the values of the coefficients of labor intensity and the cost of control and information processing are reduced by 5-6 times.

Structure and volume of the dissertation. The structure of the dissertation consists of an introduction, four chapters, a conclusion, a list of used literature and appendix. The volume of the thesis consists of 112 pages.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (Часть I; Part I)

1. Жуманов И.И., Ахатов А.Р., Тишликов С.А. Алгоритмы контроля достоверности зашифрованной текстовой информации по цифровым и числовым модулям // «Химическая технология. Контроль и управление» - ТГТУ, Ташкент, 2009- №1. - с.35-41. (05.00.00; № 12).

2. Жуманов И.И., Ахатов А.Р., Тишликов С.А. Эффективность контроля качества передачи и обработки текстовой информации на основе программной модели выявления ошибок // Журнал «Вестник ТУИТ», Ташкентский университет информационных технологий- Ташкент, 2009. - №1. - с.71-77. (05.00.00; № 31).

3. Жуманов И.И., Тишликов С.А. Алгоритмы обнаружения и исправления ошибок в текстах на основе сочетания моделей словарного и без словарного контроля // Журнал «Вестник ТУИТ», Ташкентский университет информационных технологий - Ташкент, 2011.-№2. - с.21-26.(05.00.00; № 31).

4. Жуманов И.И., Ахатов А.Р., Тишликов С.А. Адаптивные алгоритмы оптимизации процессов обнаружения и исправления ошибок в системе обработки текстов // Журнал «Вестник ТУИТ», Ташкентский университет информационных технологий - Ташкент, 2011. - №3. -с. 31-37.(05.00.00;№31).

5. Ахатов А.Р., Жуманов И.И., Тишликов С.А. Способы описания ядра функционирования программной системы контроля орфографии на основе программных моделей в среде Сфинкс // Журнал «Вестник ТУИТ», Ташкентский университет информационных технологий - Ташкент, 2012. - №2. - с.13-19. (05.00.00; № 31).

6. Тишликов С.А. Модели и алгоритмы проверки достоверности элементов текста в системе контроля орфографических ошибок // «Проблемы информатики и энергетики», Издательство «Фан» АН РУз, - Ташкент, 2012. - № 4-5. - с.91-96. (05.00.00; № 5).

7. Жуманов И.И., Тишликов С.А. Оптимизация параметров нечеткой гиперсети систем контроля и коррекции орфографии электронных документов // «Химическая технология. Контроль и управление» - ТГТУ, Ташкент, 2013. - № 6(54) . - с.69-75. (05.00.00; № 12).

8. Тишликов С.А., Ахатов А.Р. Оптимизация достоверности передачи и обработки электронных документов на основе свойств моделей переходного процесса // «Химическая технология. Контроль и управление», ТГТУ, Ташкент, 2015. - №1(61). - с.88-95. (05.00.00; № 12).

9. Ахатов А.Р., Тишликов С.А. Методы повышения достоверности передачи и обработки информации в системах электронного документооборота на основе нечеткой семантической гиперсети // Научный журнал «Проблемы вычислительной и прикладной математики», Ташкент. - №3(5), 2016. – с. 10-20. (05.00.00; № 23)

10. Jumanov I.I., Karshiev Kh.B., Tishlikov S.A. Examination of the Efficiency of Algorithms for Increasing the Reliability of Information on Criteria of Harness and the Cost of Processing Electronic Documents// International Journal of Innovative Technology and Exploring Engineering(TM) Volume-9 Issue-1, November 10, 2019. p.p. 4133- 4139.(№3;Scopus;IF=0.6).

II бўлим (Часть II; Part II)

11. Жуманов И.И., Тишликов С.А. The control of the text information transfer and processing reliability on the basis of technology of parallel computing// Proceedings international training-seminars on mathematics, SamSU and Malaysian Mathematical sciences society, Samarkand, 2011, p.p. 215-217

12. Ахатов А.Р., Тишликов С.А. Парсинговое моделирование компонентов и оценка качества системы контроля и коррекции орфографии// Журнал «Актуальные проблемы современной науки», №6 (50) 2009 г., М, изд-во «Спутник+», ISSN 1680-2721, с.180-185. (05.00.00;)

13. Ахатов А.Р., Тишликов С.А. Моделирование процессов контроля орфографии на основе парсингового представления n-граммной языковой модели // Журнал «Техника и технология», №6, 2009 г., М, изд-во «Спутник+», ISSN 1811-3532, с.43-48. (05.00.00;)

14. Жуманов И.И., Тишликов С.А. Интерактивное обнаружение и исправление ошибок в текстах систем электронного документооборота// «Илмий тадқиқотлар ахборотномаси» Илмий-назарий, услубий журнал, СамДУ, 2011, №1 (65). с.26-31 (05.00.00;)

15. Ахатов А.Р., Тишликов С.А. Оптимизация параметров гиперсети базы данных и базы знаний систем контроля и коррекции орфографии // Илмий тадқиқотлар ахборотномаси, СамДУ.– Самарканд, 2013. - № 5 (81). – с. 47-55.

16. Тишликов С.А., Ахатов А.Р. Обеспечение целостности, сохранности, достоверности информации на основе семантических свойств и структурной информационной избыточности электронных документов // Илмий тадқиқотлар ахборотномаси, СамДУ.– Самарканд, 2015. - № 5 (93). – с. 42-48.

17. Жуманов И.И., Тишликов С.А. Моделирование статистики искажений для обнаружения и исправления ошибок в системах обработки текстовой информации// Доклады респ. науч.-техн. конф. «Современное состояние и перспективы развития информационных технологий», том I, Ташкент, 2011, с.78-83.

18. Ахатов А.Р., Тишликов С.А. Алгоритмы обеспечения достоверности информации в системе электронного документооборота на основе семантической избыточности // Труды X Международной Азиатской школы-семинара «Проблемы оптимизации сложных систем», 25 июля – 5 августа 2014 г., Иссык-Кульская область, с. Булан-Соготу – Алматы, 2014 - с. 74-78

19. Akhatov.A.R, Tishlikov S. A., Isroilov N.A. Control of authenticity of information transfer and processing on the basis of mechanisms for adjustment of transients identification models parameters // In proceedings of the Ninth World

Conference on Intelligent Systems for Industrial Automation, 25-27 October, 2016. – Tashkent, Uzbekistan, 2016. – P.211-215.

20. Jumanov I.I., Tishlikov S. A. Control of integrity and authenticity of electronic documents on the basis of genetic principles of tests formation and generation // In proceedings of the Eight World Conference on Intelligent Systems for Industrial Automation, 25-27 November, 2014. – Tashkent, Uzbekistan, 2014. – P.242-246

21. Jumanov I.I., Tishlikov S. A. The system of transmitted text information control on the basis of identification of genetic algorithms // In proceedings of the Seventh World Conference on Intelligent Systems for Industrial Automation, 25-27 November, 2012. – Tashkent, Uzbekistan, 2012. – P.196-200

22. Jumanov I.I., Tishlikov S. A. Increase of efficiency of word-forms search and processing for the control and correction of spelling mistakes in the electronic texts // The International Scientific Conference «Modern problems of applied mathematics and information technologies – Al-Khorezmiy 2012», Volume № 2, 19-22 December, National University of Uzbekistan. - Tashkent, 2012. – p.p. 93-96

23. Тишликов С.А. Контроль достоверности документов систем автоматизированного делопроизводства на основе структурно-технологической избыточности // Материалы X-й Международной научно-практической конференции «Современные инструментальные системы, информационные технологии и инновации», 19-23 марта 2013 г., с. 247-252.

24. Тишликов С.А. Оценка эффективности методов внеконтекстного контроля достоверности текстовой информации // Материалы VIII международной научно-практической конференции «Научная дискуссия: вопросы технических наук», (4 апреля 2013 г.) – М, «Международный центр науки и образования», 2013. – с. 58-63

25. Jumanov I.I., Tishlikov S. A. Method of stochastic search in the system to monitoring and correcting of spelling in electronic texts // 2013 International Conference in Central asia on Internet (ICI), Tashkent, 8-10 october 2013, Section 7, IEEE. – Tashkent, 2013.

26. Жуманов И.И., Тишликов С.А. Оптимизация параметров нечеткой гиперсети систем контроля и коррекции орфографии электронных документов // Научная дискуссия: вопросы технических наук, № 11 (13): М., Изд. «Международный центр науки и образования», 2013. – с. 18-23

27. Жуманов И.И., Тишликов С.А. Оптимизация контроля достоверности электронных документов на основе моделей переходного процесса и рейтинговых оценок // Доклады Республиканской научно-технической конференции «Современное состояние и перспективы применения информационных технологий в управлении», ТУИТ, 7-8 сентября 2015 г, с. 141-147

28. Тишликов С.А. Ахборот ишончилигини луғатсиз ортиқчаликка эга кодлаштириш ва модулли йиғиндилар асосида ошириш алгоритмлари // Международная конференция «Наука и инновации»: сборник научных

трудов: 26 ноября, 2020 года / Министерство инновационного развития Республики Узбекистан, Центр передовых технологий. – Ташкент: Центр передовых технологий, 2020 г, б. 517-520

29. Тишликов С.А. Программные комплексы повышения достоверности передачи и оптимизации обработки информации документов делопроизводства в низовых системах «Электронного правительства» // Свидетельство об официальной регистрации программы для ЭВМ, Агентство по интеллектуальной собственности при министерстве юстиции РУз, №DGU 10447, 10.03.2021

30. Тишликов С.А., Ахатов А.Р., Жуманов И.И. Программа обнаружения и исправления ошибок в электронных текстах на узбекском языке на основе стохастического поиска словоформ // Свидетельство об официальной регистрации программы для ЭВМ, Агентство по интеллектуальной собственности РУз, №DGU 02693, 24.01.2013

31. Жуманов И.И., Ахатов А.Р., Тишликов С.А. Ўзбек тилидаги матнлар таркибидаги имло ҳатоларни сўзларнинг хеш-кодлари жадвалларини қўллаш асосида назорат қилиш дастури // Государственное патентное ведомство РУз, Свидетельство № DGU 01840, 07.10.2009.

32. Жуманов И.И., Ахатов А.Р., Тишликов С.А. Ўзбек тилидаги матнлар таркибидаги имло ҳатоларни морфемлар таҳлили ва синтези асосида назорат ва таҳрир қилиш дастурий тизим интерфейси // Государственное патентное ведомство РУз, Свидетельство № DGU 01842, 07.10.2009.