

ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՅՄԱՆ ՊՐՈՔԼԵՄՆԵՐԻ
ԻՆՍՏԻՏՈՒՏ

ՎԱՐԱԶԴԱՏ ԿՈՐՅՈՒՆԻ ԱՎԵՏԻՍՅԱՆ

**ԳԻՏԵԼԻՔՆԵՐԻ ԱՏՈՒԳՄԱՆ ԹԵՍԵՐԻ ՈՐԱԿԻ ԳՆԱՀԱՏՄԱՆ ՄԵԹՈՂԻ ԵՎ
ԾՐԱԳՐԱՅԻՆ ՀԱՄԱԿԱՐԳԻ ՄՇԱԿՈՒՄ**

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

ՍԵՂՄԱԳԻՐ

Երևան – 2016

ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ НАН РА

ВАРАЗДАТ КОРЮНОВИЧ АВЕТИСЯН

**РАЗРАБОТКА МЕТОДА И ПРОГРАММНОЙ СИСТЕМЫ ОЦЕНКИ КАЧЕСТВА
ТЕСТОВ ДЛЯ ПРОВЕРКИ ЗНАНИЙ**

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата технических наук по специальности

05.13.04 – “Математическое и программное обеспечение вычислительных машин, комплексов, систем и сетей”

Ереван – 2016

Ատենախոսության թեման հաստատվել է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում

Գիտական ղեկավար՝	Ֆիզ.մաթ.գիտ. դոկտոր	Մ. Ե. Հարությունյան
Պաշտոնական ընդդիմախոսներ՝	Ֆիզ.մաթ.գիտ. դոկտոր	Հ. Բ. Մարանջյան
	Ֆիզ.մաթ.գիտ. թեկնածու	Գ. Ս. Չիլինգարյան

Առաջատար կազմակերպություն՝ Հայաստանի ազգային պոլիտեխնիկական համասարան

Պաշտպանությունը կայանալու է 2016թ. հունիսի 10-ին, ժ. 16:00-ին ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում գործող 037 «Ինֆորմատիկա և հաշվողական համակարգեր» մասնագիտական խորհրդի նիստում հետևյալ հասցեով՝ Երևան, 0014, Պ. Սևակի 1:

Ատենախոսությանը կարելի է ծանոթանալ ինստիտուտի գրադարանում:
Սեղմագիրն առաքված է 2016թ. մայիսի 10-ին:

Մասնագիտական խորհրդի գիտական
քարտուղար, ֆիզ.մաթ.գիտ.դոկտոր



Հ. Գ. Սարգսյան

Тема диссертации утверждена в Институте проблем информатики и автоматизации НАН РА

Научный руководитель:	доктор физ.-мат. наук	М. Е. Арутюнян
Официальные оппоненты:	доктор физ.-мат. наук	Г. Б. Маранджян
	канд физ.-мат. наук	Г. С. Чилингарян

Ведущая организация: Национальный Политехнический Университет
Армении

Защита состоится 10-го июня 2016г. в 16:00 на заседании специализированного совета 037 “Информатика и вычислительные системы” в Институте проблем информатики и автоматизации НАН РА по адресу: 0014, г. Ереван, ул. П. Севака 1.

С диссертацией можно ознакомиться в библиотеке ИПИА НАН РА.
Автореферат разослан 10-го мая 2016г.

Ученый секретарь специализированного
Совета, доктор физ.мат.наук



А. Г. Саруханян

ԱՇԽԱՏԱՆՔԻ ԸՆԴՀԱՆՈՒՐ ԲՆՈՒԹԱԳԻՐԸ

Թեմայի արդիականությունը

Համաշխարհային փորձը վկայում է գիտելիքի ստուգման համար կիրառվող քննական թեստերի արդյունավետության մասին: Թեստերը օգտագործվում են (ինքնա)ուսուցման և (ինքնա)ստուգման համար, թույլ են տալիս ստանալ գիտելիքների մակարդակի բավականին օբյեկտիվ գնահատական, բացահայտել ուսուցման բացթողումները: Ուսումնական հաստատություններում մանկավարժական թեստերը օգտագործվում են ընդունելության քննությունների, գիտելիքների ընթացիկ ստուգման համար: Հեռուսուցման ժամանակ լայն կիրառվում են թեստավորման համակարգչային տեխնոլոգիաները:

Համակարգչային թեստավորման ձևով քննությունը թույլ է տալիս ստուգել գիտելիքներն առարկայի նյութի մեծ մասով, ապահովում է գնահատականի օբյեկտիվությունը, ազատում է դասավանդողին գրավոր աշխատանքների աշխատատար ստուգումներից:

Կարևոր է նշել, որ թեստավորման և գիտելիքների գնահատման օբյեկտիվությունը մեծապես կախված է թեստերի բովանդակությունից և որակից: Ներկայումս շատ դեպքերում կիրառվում են թեստեր, որոնք չեն ենթարկվում վերլուծության, որակի գնահատման և փորձաքննության, բնութագրվում են հուսալիության և վավերության ցածր, որոշ դեպքերում՝ անբավարար ցուցանիշներով, որի արդյունքում էլ չի ապահովվում գիտելիքների ստուգման օբյեկտիվությունը, արդյունքների համապատասխանությունը: Այդ իսկ պատճառով թեստերի և թեստային առաջադրանքների որակական բնութագրերի վերլուծության և գնահատման նոր արդյունավետ գործիքների մշակումը արդիական գիտական խնդիր է:

Վերջին տարիներին Հայաստանի Հանրապետությունում ևս լայն կիրառություն են գտնում թեստավորման տեխնոլոգիաները: Դպրոցական ավարտական և ԲՈՒՀ-երի ընդունելության միասնական քննությունները կազմակերպվում են թեստավորման եղանակով՝ ՀՀ կառավարության աշխատակազմի «Գնահատման և թեստավորման կենտրոն» ՊՈԱԿ-ի կողմից: Համակարգչային թեստավորման եղանակով քննություններ են կազմակերպվում ԲՈՒՀ-երում, ինչպես նաև հետբուհական կրթության ոլորտում, օրինակ ասպիրանտուրայի ընդունելության Ինֆորմատիկայի ստուգաբլքը, անգլերենի «TOEFL» քննությունը, որակավորման Ինֆորմատիկայի ստուգաբլքը:

Թեստի որակի գնահատումը կարող է իրականացվել վիճակագրական եղանակով, փորձնական թեստավորման միջոցով ստացված արդյունքների վերլուծության հիման վրա: Վիճակագրական եղանակով թեստի որակը գնահատելու

համար գոյություն ունեն տեսություններ, որոնցից առավել լայն կիրառություն են գտել թեստերի դասական տեսությունը¹ (Classical Test Theory-CTT) և թեստերի ժամանակակից մաթեմատիկական տեսությունը² (Item Resonse Theory-IRT): Երկու տեսություններում էլ սահմանվում են թեստի և թեստային առաջադրանքների որակական ցուցանիշները, դրանց հաշվարկման, գնահատման մեթոդները և մաթեմատիկական վիճակագրական ապարատը: Հարկ է նշել, որ վիճակագրական ապարատը բավականին բարդ է տարբեր բնագավառների մանկավարժների համար, իսկ հաշվարկները աշխատատար են նեղ մասնագետի համար: Այդ խնդիրը լուծելու համար մշակվում են համակարգեր, ծրագրային փաթեթներ, որոնց միջոցով իրականացվում է փորձնական թեստավորման արդյունքների վերլուծություն և թեստի որակական բնութագրերի գնահատում:

Ներկայումս գոյություն ունեն թեստերի որակական վերլուծության բազմաթիվ ծրագրային փաթեթներ և համակարգեր³, որոնցից առավել լայնորեն կիրառվում են՝ CITAS, Iteman 4, Xcalibre 4, Winsteps, Facets, jMetrik, RUMM 2030, ACER ConQuest, IRTPRO, ConstructMap, Bigsteps, GENIRV, RESCEN: Թվարկվածներից յուրաքանչյուրն ունի իրեն բնորոշ հատկանիշներ, հնարավորություններ, առավելություններ, թերություններ և նախապատվություն՝ կախված կիրառական նշանակությունից, ոլորտից և բնույթից: Հիմնական թերություններից կարելի է առանձնացնել այն, որ նմանատիպ համակարգերը չունեն օգտվողին կողմնորոշված պարզ ինտերֆեյս, երբեմն աշխատելու համար պահանջվում են տեխնիկական խորը գիտելիքներ, հրամանային տողով աշխատելու հմտություններ: IRT –ով վերլուծություն անցկացնող ծրագրերը բազմագործառության են և կիրառվում են տարբեր բնագավառներում չափումներ գնահատելու համար, իսկ միայն թեստավորման վերաբերյալ վերլուծություն անցկացնելու համար անհրաժեշտ է ծրագրից գտնել, առանձնացնել և խմբավորել միայն թեստավորման մոդուլները, որն այդքան էլ դյուրին գործ չէ: Թերություն կարելի է համարել նաև այն, որ թեստավորման արդյունքների վերլուծությունը հիմնականում ստացվում է բազմաթիվ աղյուսակների տեսքով, որոնք պահպանվում են TXT ձևաչափով: Գրաֆիկներն էլ իրենց հերթին ստացվում են առանձին ֆայլերի տեսքով՝ JPG կամ PNG ձևաչափով: Այսինքն մեկ միասնական ֆայլի տեսքով հաշվետվություն ստանալու համար անհրաժեշտ է լինում տարբեր ֆայլերից կատարել խմբագրումներ և ստանալ նոր հաշվետվություն, ինչը հարմար և կիրառելի չէ մանկավարժի համար:

¹ Crocker L., Algina J., Introduction to Classical and Modern Test Theory, New York, Wadsworth Pub Co, 2006, 527 p.

² DeMars C., Item Response Theory, New York, Oxford University Press, 2010, 144 p.

³ Rasch Measurement Analysis Software Directory, <http://www.rasch.org/software.htm>

Բացի այդ, գոյություն ունեցող ծրագրերը հիմնականում անգլերեն լեզվով են, վճարովի են և կապված բազմագործառնության հետ ունեն բարձր արժեքներ:

Այս իրավիճակում արդիական էր մշակել թեստերի որակի գնահատման ավելի պարզ մեթոդ և հայերեն լեզվով ծրագրային համակարգ:

Հետազոտության նպատակը

Թեստերի որակի ստուգման, գնահատման գործընթացի կազմակերպման համար խնդիր էր դրվել մշակել նոր եղանակ և ծրագրային իրականացում, որը կապահովի բազմակողմանի վերլուծություն հայերեն լեզվի սատարմամբ:

Հետազոտման օբյեկտը

Ուսումնասիրության օբյեկտներն են թեստերի որակի վերլուծության և գնահատման համակարգերն իրականացնող տեխնոլոգիաները, մասնավորապես, Web Server-ի վրա աշխատանքի կազմակերպման սկզբունքները, Java ծրագրավորման լեզվով Microsoft Office փաստաթղթերի ստեղծման և ձևավորման ֆունկցիաները, Java և PHP ծրագրավորման լեզուներով CSV, TXT ֆայլերի հետ աշխատելու միջոցները, PHP ծրագրավորման լեզվով արտաքին ծրագրերի գործարկման հնարավորությունները:

Հետազոտման մեթոդները

Ատենախոսության մեջ կիրառվում են վիճակագրական ու ինֆորմացիայի տեսության մեթոդներ, CTT և IRT տեսությունների սկզբունքներ: Օգտագործվում են Java psychometrics բաց ծրագրային կոդով գրադարանի⁴ CTT, IRT ցուցանիշների հաշվարկման մաթեմատիկական և վիճակագրական մեթոդները:

Հետազոտության գիտական նորույթը

1. Մշակվել է թեստերի որակի գնահատման նոր մոտեցում՝ հիմնված ինֆորմացիայի տեսության⁵ տարրերի վրա, որն ապահովում է ավելի արագ հաշվարկներ և վավեր վերլուծության իրականացման համար պահանջում է շատ ավելի քիչ թեստավորման արդյունքներ:
2. Մշակված թեստերի որակի գնահատման հայալեզու սատարմամբ ծրագրային համակարգը գործարկվում է սերվերի վրա և ունի web հենքով օգտվողի ինտերֆեյս, ինչն էլ թույլ է տալիս դիմել համակարգին լոկալ ցանցի կամ համացանցի միջոցով՝ գործարկելով այն web զննարկիչի միջավայրում:

⁴ Java library for psychometric methods, <https://www.java.net/projects/psychometrics>.

⁵Cover T. M., Thomas J. A., Elements of Information Theory, 2nd Edition, New York, Wiley, 2006, 776p.

3. Համակարգի առանձին բաղադրիչ հանդիսացող վիճակագրական և մաթեմատիկական ցուցանիշները հաշվարկող մոդուլը կարող է կիրառվել նաև առանձին համակարգչի վրա որպես desktop հավելված:
4. Մշակված համակարգը կախված չէ սերվերի օպերացիոն համակարգի տեսակից:

Ստացված արդյունքների կիրառական նշանակությունը

Մշակված պարզ և մատչելի web ինտերֆեյսով համակարգն օգտվողներին հնարավորություն է տալիս առանց տիրապետելու թեստերի վիճակագրական տեսության բարդ ապարատին և առանց տեղադրելու լրացուցիչ ծրագրային միջոցներ, կազմակերպել մշակված թեստերի որակի վերլուծություն և գնահատում:

Համակարգի միջոցով գեներացվում է թեստերի որակի վերլուծության և գնահատման հաշվետվություն, որը խմբագրվող rtf (rich text format) ձևաչափի ֆայլ է և բովանդակում է ինչպես հաշվարկվող և գնահատվող ցուցանիշները, աղյուսակները, գրաֆիկները, այնպես էլ նկարագրություններ և ցուցումներ դրանց վերաբերյալ:

Համակարգը կարող է կիրառվել թեստերի մշակմամբ զբաղվող տարբեր մասնագիտությունների մանկավարժների և թեստավորման անցկացմամբ զբաղվող կազմակերպությունների (ուսումնական հաստատություններ, թեստավորման կենտրոններ և այլն) կողմից՝ թեստավորման արդյունքների վերլուծության և թեստերի որակի գնահատման նպատակով:

Ներդրումներ

Մշակված համակարգը ներդրվել է Գավառի պետական համալսարանի կողմից իրականացվող համակարգչային թեստավորումով քննական գործընթացի մեջ: Ներդրված համակարգի միջոցով վերլուծվում են թեստավորման արդյունքները և գնահատվում է թեստերի որակը:

Պաշտպանության ներկայացվող հիմնական դրույթները

- Առաջարկվել է թեստերի որակի գնահատման նոր մեթոդ՝ հիմնված ինֆորմացիայի տեսության տարրերի վրա, որն ապահովում է ավելի արագ հաշվարկներ:
- Որոշվել են առաջարկված մեթոդի ցուցանիշների ընդունելի սահմանները:
- Գնահատվել է վավեր վերլուծության համար անհրաժեշտ թեստավորման արդյունքների քանակը, որը էականորեն ավելի փոքր է մյուս մեթոդների հետ համեմատած:
- Մշակվել է CTT, IRT տեսությունների և առաջարկված նոր մեթոդի հիման վրա թեստերի որակի վերլուծություն և գնահատում իրականացնող երկվեզու ծրագրային համակարգ, որը գործարկվում է սերվերի վրա և ունի web

հենքով օգտվողի ինտերֆեյս, ինչն էլ թույլ է տալիս դիմել համակարգին լոկալ ցանցի կամ համացանցի միջոցով՝ գործարկելով այն web զննարկիչի միջավայրում: Համակարգի առանձին բաղադրիչ հանդիսացող վիճակագրական և մաթեմատիկական ցուցանիշները հաշվարկող մոդուլը կարող է կիրառվել նաև առանձին համակարգչի վրա որպես desktop հավելված: Մշակված համակարգը կախված չէ սերվերի օպերացիոն համակարգի տեսակից:

Ապրոբացիա

Ատենախոսության արդյունքները զեկուցվել են՝

- ITA 2015 – ITHEA ISS Joint International Events on Informatics Summer Session, Varna, Bulgaria, June 29 - July 12, 2015,
- «Համակարգչային գիտություններ և տեղեկատվական տեխնոլոգիաներ (CSIT 2015)» 10-րդ գիտաժողովում, Երևան, սեպտ. 28- հոկտ. 2, 2015,
- «Բնագիտական խնդիրներ հետազոտական էլեկտրոնային ենթակառուցվածքի միջավայրում» գիտաժողովում, Երևան, դեկտեմբերի 18-20, 2015,
- Գնահատման և թեստավորման կենտրոնի համակարգչային և վերլուծական աշխատանքների բաժնում,
- ինչպես նաև ՀՀ ԳԱԱ ԻԱՊԻ ընդհանուր սեմինարում:

Հրապարակումներ

Ատենախոսության հիմնական արդյունքները հրապարակվել են հինգ գիտական հոդվածներում, որոնց ցանկը բերված է սեղմագրի վերջում [1-5]:

Ատենախոսության կառուցվածքը և ծավալը

Ատենախոսությունն իր մեջ ներառում է ներածություն, հինգ գլուխ, եզրակացություններ, գրականության ցանկ՝ իր 92 հղումներով: Ատենախոսության ընդհանուր ծավալը՝ 113 էջ՝ 13 աղյուսակներով և 29 նկարներով:

ԱՇԽԱՏԱՆՔԻ ՀԻՄՆԱԿԱՆ ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆ

Ներածության մեջ հիմնավորվել է ատենախոսության թեմայի արդիականությունը, ներկայացվել են հետազոտության նպատակն ու խնդիրները, գիտական նորոյթը, պաշտպանությանը ներկայացվող հիմնական դրույթները, հետազոտության տեսական և գործնական նշանակությունը:

Ատենախոսության **առաջին՝ «Թեստերի որակի վերլուծության**

տեսությունները» խորագրով գլուխը բաղկացած է չորս բաժիններից և նվիրված է վիճակագրական եղանակով թեստերի ու դրանց որակի գնահատման համար նախատեսված CTT և IRT տեսությունների նկարագրությանը:

Բաժին 1.1-ում նկարագրվել են գիտելիքների ստուգման թեստերի տեսակները, թեստային առաջադրանքների ձևերը, թեստավորման արդյունքների վերլուծման սկզբունքները: Մանրամասն ներկայացվել են փակ, բաց, համապատասխանության հաստատման, ճիշտ հաջորդականության որոշման տիպի առաջադրանքների և այդ առաջադրանքների հիման վրա մշակվող թեստավորման արդյունքների մատրիցաների ձևավորման առանձնահատկությունները:

Բաժին 1.2-ում ներկայացվել է CTT տեսությունը, նկարագրվել են այս տեսության շրջանակներում գնահատվող թեստի որակական բնութագրերը, մասնավորապես՝ հուսալիությունը, վավերությունը, բարդությունը, տարբերակիչ ունակությունը, ստանդարտ շեղումը: Մանրամասն ներկայացվել են նկարագրված բնութագրերի հաշվարկման մեթոդները, արժեքների և համապատասխան գնահատման աղյուսակները: Նկարագրվել են հուսալիության մի շարք գործակիցներ, ինչպես նաև հուսալիության հաշվարկման տարբեր մեթոդներ: Ներկայացվել են առաջադրանքների միջև կոռելյացիայի և առաջադրանք-անհատական միավոր կոռելյացիայի գործակիցների նշանակությունը և դրանց հաշվարկները:

Բաժին 1.3-ում դիտարկվում է թեստերի ժամանակակից մաթեմատիկական տեսությունը, դրա շրջանակներում կիրառվող թեստերի որակի գնահատման մաթեմատիկական մոդելները: IRT տեսությունն ուղղված է փորձարկվողի և առաջադրանքների պարամետրերի լատենտային որակների գնահատմանը: Այս տեսության շրջանակներում առաջադրվում է հետևյալ խնդիրը՝ ըստ թեստին տված փորձարկվողի պատասխանների, գնահատել առաջադրանքների բարդության և փորձարկվողի պատրաստվածության մակարդակի լատենտային պարամետրերի արժեքները: Մանրամասն ներկայացվել են միաճյուղ գնահատման երեք մոդելներ՝ Ռաշի մեկ պարամետրով (1PLM), Բիրնբաումի երկու պարամետրով (2PLM) և Բիրնբաումի երեք պարամետրով (3PLM) մոդելները, բազմաճյուղ գնահատման երկու մոդելները՝ մասնակի միավորներով (PCM) և ընդհանրացված մասնակի միավորներով (GPCM): Նկարագրվել են այս տեսության շրջանակներում կիրառվող թեստերի և թեստային առաջադրանքների բնութագրական ֆունկցիաները և դրանց գրաֆիկների կառուցման սկզբունքները: Ներկայացվել են առաջադրանքի բարդության «b պարամետր»-ի, տարբերակիչ ունակության «a պարամետր»-ի և պատահականության գույակելու «c պարամետր»-ի նկարագրությունները, այդ պարամետրերի հիման վրա առաջադրանքների բնութագրական ֆունկցիաների գրաֆիկների կառուցման սկզբունքները: Նկարագրվել են գնահատվող պարամետրերի արժեքների չափումների

որակը բնութագրող առաջադրանքների (IIF) և թեստի ինֆորմացիոն (TIF) ֆունկցիաները, վերջիններիս որոշումը ըստ մոդելի, ինչպես նաև այդ ֆունկցիաների և չափումների ստանդարտ սխալի ֆունկցիայի (SEM) միջև կախվածությունը:

Բաժին 1.4-ը նվիրված է վերոնշյալ տեսությունների համեմատական վերլուծությանը, առավելությունների և թերությունների նկարագրությանը: CTT-ի առավելություններից կարելի է նշել մաթեմատիկական պարզ ապարատը՝ IRT-ի համեմատությամբ, վերլուծության համար անհրաժեշտ փորձարկվողների համեմատաբար փոքր բավարար քանակը: IRT-ի հիմնական առավելություններից կարելի է նշել այն, որ ի տարբերություն CTT-ի, թեստային առաջադրանքի բարդության գնահատականը կախված չէ փորձարկվողների խմբի ընտրությունից, փորձարկվողների պատրաստվածության մակարդակի գնահատականն էլ կախված չէ կիրառվող թեստային առաջադրանքների բանկից, ինչպես նաև առաջադրանքների և փորձարկվողների պարամետրերը չափվում են նույն սանդղակով:

Ատենախոսության **երկրորդ գլուխը՝ «Առկա համակարգերի հետազոտություն»** վերնագրով, նվիրված է առավել հայտնի թեստերի որակական վերլուծության անցկացման ծրագրերի, համակարգերի ուսումնասիրությանը, առավելությունների ու թերությունների նկարագրությանը:

Բաժին 2.1-ում ուսումնասիրվել են հետևյալ ծրագրերը և համակարգերը՝ CITAS, Itean 4, Xcalibre 4, Winsteps, Facets, jMetrik, RUMM 2030, ACER ConQuest, IRTPRO, ConstructMap, Bigsteps, GENIRV, RESCEN: Ներկայացվել են այդ համակարգերի հիմնական բնութագրերը և առանձնահատկությունները: Թեստերի որակական վերլուծության անցկացման համակարգերի հնարավորությունները գնահատելիս առանձնացվում են հետևյալ հիմնական բնութագրերը՝ վերլուծության համար նախատեսված թեստավորման արդյունքների մուտքագրման ձևաչափը, թեստի CTT և IRT որակական բնութագրերի բազմազանությունը, օգտվողի ինտերֆեյսը և դրա մատչելիությունը, վերլուծության ենթարկվող թեստի հնարավոր փորձարկվողների և առաջադրանքների քանակը, վերլուծության արդյունքում ստացվող աղյուսակների, գրաֆիկների և դրանց ուղեկցող նկարագրությունների, հաշվետվությունների ստացման ձևաչափերը, որպես ծրագիր՝ հարթակի (պլատֆորմի) վրա գործարկվելու առանձնահատկությունները:

Բաժին 2.2-ում կատարված է համեմատական վերլուծություն, նկարագրված են կարևոր առանձնահատկությունները: Կարևոր է նշել, որ գոյություն ունեցող համակարգերը չունեն օգտվողին կողմնորոշված պարզ ինտերֆեյս, երբեմն պահանջում են հրամանային տողով աշխատելու հմտություններ, որոշ դեպքերում նաև տեխնիկական խորը գիտելիքներ: IRT –ով վերլուծություն անցկացնող ծրագրերը բազմագործառության են և կիրառվում են տարբեր բնագավառներում չափումներ

գնահատելու համար, իսկ միայն թեստավորման վերաբերյալ վերլուծություն անցկացնելու համար անհրաժեշտ է ծրագրից գտնել, առնաձնացնել և խմբավորել միայն թեստավորման մոդուլները, որը այդքան էլ դյուրին գործ չէ: Նշված է նաև այն, որ թեստավորման արդյունքների վերլուծությունը հիմնականում ստացվում է բազմաթիվ աղյուսակների տեսքով, որոնք պահպանվում են TXT ձևաչափով: Գրաֆիկներն էլ իրենց հերթին ստացվում են առանձին ֆայլերի տեսքով՝ JPG կամ PNG ձևաչափով: Այսինքն մեկ միասնական ֆայլի տեսքով հաշվետվություն ստանալու համար անհրաժեշտ է լինում տարբեր ֆայլերից կատարել խմբագրումներ և ստանալ նոր հաշվետվություն, ինչը հարմար և կիրառելի չէ մանկավարժի համար: Բացի այդ ուսումնասիրված համակարգերը հիմնականում անգլերեն լեզվով են, վճարովի են և կապված բազմագործառնության հետ ունեն բարձր արժեքներ: Հիմնավորվել է հայերեն լեզվի սատարումով և միաժամանակ IRT, CTT տեսությունների հիման վրա թեստերի վերլուծություն և որակի գնահատում իրականացնող, օգտվողի պարզ և մատչելի web ինտերֆեյս ունեցող նոր համակարգ մշակելու անհրաժեշտությունը:

Երրորդ գլուխը՝ վերնագրված է «Թեստերի որակի գնահատման նոր մոդելի մշակում» և նվիրված է նոր մոտեցման նկարագրությանը: Առաջարկված գնահատման նոր մոդելը հիմնված է Շենոնի ինֆորմացիայի չափի մեծությունների վրա:

Բաժին 3.1-ում նկարագրված են մոդելի մեջ կիրառվող Շենոնի էնտրոպիան, պայմանական էնտրոպիան, փոխադարձ միջին ինֆորմացիան: Թեստային առաջադրանքը դիտարկվում է որպես պատահական X մեծություն, որը կարող է ընդունել երկու արժեք՝ 1 - ճիշտ պատասխանելու դեպքում և 0 - սխալ պատասխանելու դեպքում՝ համապատասխանաբար p և $1-p$ հաճախականություններով, որտեղ p -ն թեստին տրված ճիշտ պատասխանների քանակի հարաբերությունն է փորձարկվողների ընդհանուր քանակին: Թեստի որակի գնահատման նոր մոդելը ձևակերպվել է էնտրոպիայի և միջին փոխադարձ ինֆորմացիայի հետևյալ հատկությունների հիման վրա:

1. Երկուական տվյալների դեպքում (0 կամ 1) տեղի ունեն հետևյալ անհավասարությունները՝

$$0 \leq H(X) \leq 1,$$

ընդ որում հավասարությունը ձախից տեղի ունի այն և միայն այն դեպքում, երբ $p_i = 0$ կամ $p_i = 1$, իսկ աջից հավասարությունը տեղի ունի այն և միայն այն դեպքում, երբ $p_i = 1/2$:

2. Տեղի ունեն հետևյալ անհավասարությունները՝

$$0 \leq I(X_i \wedge X_j) \leq \min[H(X_i), H(X_j)], \quad i, j = \overline{1, N}$$

ընդ որում հավասարությունը ձախից տեղի ունի այն և միայն այն դեպքում, երբ X_i և X_j մեծությունները միմյանցից անկախ են:

3. Պայմանական էնտրոպիան չի գերազանցում ոչ պայմանականին՝

$$H(X_i|X_j) \leq H(X_i), \quad i, j = \overline{1, N}$$

ընդ որում հավասարությունը տեղի ունի այն և միայն այն դեպքում, երբ X_i և X_j մեծությունները միմյանցից անկախ են:

Բաժին 3.2-ում նկարագրված է առաջարկված թեստի որակի գնահատման նոր մոդելը, բերված է մոդելի մաթեմատիկական ապարատի նկարագրությունը: Ներկայացվել է նաև փորձնական թեստավորման արդյունքների հիման վրա հաշվարկների օրինակներ: Նոր մոդելում թեստի որակի գնահատման համար ձևակերպված են հետևյալ մեթոդները:

Մեթոդ 1. Եթե $H(X_i)$ մեծության արժեքը մոտ է 0-ին, ապա թեստային առաջադրանքը վատն է, այսինքն՝ կամ շատ բարդ է կամ շատ հեշտ: Եթե $H(X_i)$ մեծության արժեքը մոտ է 1-ին, ապա թեստային առաջադրանքը լավն է, այսինքն՝ ունի միջին բարդություն:

Մեթոդ 2. Եթե $I(X_i \wedge X_j)$ մեծության արժեքը մոտ է 0-ին, ապա X_i և X_j թեստային առաջադրանքները միմյանցից անկախ են: Իսկ եթե արժեքը մոտ է $\min[H(X_i), H(X_j)]$ -ին, ապա X_i և X_j թեստային առաջադրանքները կրկնում են մեկը մյուսին և փոքր էնտրոպիա ունեցողը պետք է հեռացվի թեստից:

Մեթոդ 3. $H(X_i|X_j)$ պայմանական էնտրոպիայի արժեքը ցանկալի է մոտ լինի $H(X_i)$ -ին, այսինքն X_j առաջադրանքը չնչին ազդեցություն ունենա X_i -ի անորոշության վրա:

Բաժին 3.3-ում ամփոփված են նոր մոտեցման առանձնահատկությունները:

«Թեստերի որակի գնահատման գործիքային ծրագրային համակարգի մշակում» վերնագրով չորրորդ գլխում նկարագրված է մշակված համակարգը, ներկայացված են առանձնահատկությունները, մշակման ժամանակ իրականացված տեխնիկական լուծումները:

Բաժին 4.1-ում ներկայացված են համակարգի հիմնական հնարավորությունները: Համակարգը հնարավորություն է տալիս իրականացնել թեստերի որակի վերլուծություն և գնահատում թեստերի երկու՝ CTT և IRT տեսությունների և Շենոնի ինֆորմացիայի տեսության հիման վրա: Համակարգը սատարում է IRT միաճյուղ գնահատման երեք մոդել և բազմաճյուղ գնահատման երկու մոդել: Գներազգայում են թեստի և թեստային առաջադրանքների մի շարք IRT բնութագրական ֆունկցիաների գրաֆիկներ:

Բաժին 4.2-ում նկարագրվել է համակարգի կառուցվածքը: Համակարգը բաղկացած է երեք հիմնական մոդուլից.

- օգտվողի ինտերֆեյս,
- թեստի որակի վերլուծության և գնահատման մոդուլ,

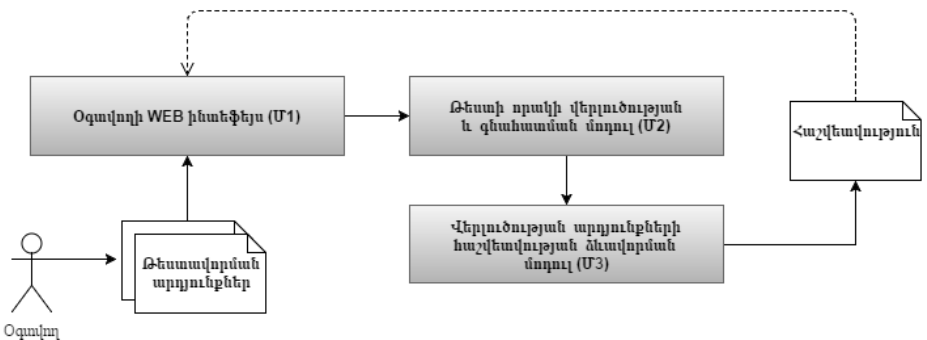
- արդյունքների մշակման և հաշվետվության ձևավորման մոդուլ:

Օգտվողի web ինտերֆեյսը նախատեսված է օգտվողի հետ փոխգործակցության ապահովման համար, մասնավորապես, այս մոդուլի միջոցով մուտքագրվում են վերլուծության ենթարկվող թեստավորման արդյունքների ֆայլերը:

Թեստի որակի վերլուծության և գնահատման մոդուլի միջոցով օգտվողի web ինտերֆեյսից ստացված տվյալների հիման վրա իրականացվում է թեստերի որակական ցուցանիշների հաշվարկ և գնահատում, համապատասխան արդյունքների ստեղծում և գրանցում տվյալների բազայում:

Ստեղծված տվյալների բազայի հիման վրա վերլուծության արդյունքների հաշվետվության ձևավորման մոդուլի միջոցով գեներացվում են հաշվետվության ֆայլերը, որոնք ներբեռնելու տարբերակով հասանելի են դառնում օգտվողի ինտերֆեյսում:

Համակարգի կառուցվածքի և գործառնության ընդհանրացված սխեման բերված է նկար 1-ում:



Նկար 1. Համակարգի կառուցվածքի և գործառնության ընդհանրացված սխեման:

Վերլուծության համար անհրաժեշտ թեստավորման արդյունքները, որոնք հանդիսանում են մուտքային ինֆորմացիա, մուտքագրվում են TXT կամ CSV ձևաչափի ֆայլերի տեսքով: Թեստի որակական վերլուծության և գնահատման արդյունքները՝ որպես ելքային ինֆորմացիա, ստացվում են մեկ ամբողջական հաշվետվության՝ rtf ձևաչափի խմբագրվող ֆայլի տեսքով: Բացի այդ առանձին CSV ձևաչափով ֆայլի տեսքով ստացվում են հաշվարկման արդյունքների աղյուսակները և գրաֆիկները նկարները:

Բաժին 4.3-ում ներկայացվել է համակարգի օգտվողի ինտերֆեյսը: Օգտվողի ինտերֆեյսը web հենքով է և գործարկվում է սերվերի վրա, ինչը թույլ է տալիս

օգտվողին դիմել համակարգին լոկալ ցանցի միջոցով կամ համացանցի միջոցով գործարկելով այն web դիտարկիչի միջավայրում: Օգտվողի ինտերֆեյսը հասանելի է երկու լեզուներով՝ հայերեն և անգլերեն, ներառված են հետևյալ բաժինները՝ «Թեստի վերլուծություն», «Որակական բնութագրեր», «Օրինակներ», «Օգնություն», «Լեզվի ընտրություն»: Օգտվողի ինտերֆեյսի գլխավոր էջը պատկերված է նկար 2-ում:

The screenshot shows a web interface for test analysis. At the top, there are navigation links in Armenian: 'ԹԵՄՏԻ ՎԵՐԼՈՒԾՈՒԹՅՈՒՆ', 'ՈՐԱԿԱԿԱՆ ԲՆՈՒԹԱԳՐԵՐ', 'ՕՐԻՆԱԿՆԵՐ', and 'ՕԳՆՈՒԹՅՈՒՆ'. There are also language selection buttons for 'ՀՅԵ' and 'ԱՆԳ'. The main heading is 'Թեստի վերլուծություն'. Below it, there are four main sections:

- Թեստի արդյունքների ֆայլ** (Test Results File) with a button labeled 'ՎԵՐԲԵՈՒՆԵԼ'.
- Թեստի պատասխանների ֆայլ** (Test Answers File) with a button labeled 'ՎԵՐԲԵՈՒՆԵԼ'.
- Թեստի վերնագիր** (Test Title) with a button labeled 'Թեստի վերնագիր'.
- Թեստի գնահատման տեսակը** (Test Evaluation Type) with radio buttons for:
 - Միանյութ (Ֆինալ կամ սարքերակի ընտրությունը)
 - Բազմանյութ (Polytomous)
 - Խառը
- IRT մոդելներ** (IRT Models) with radio buttons for:
 - 1 պարամետրով միանյութ մոդել (Ռաչի մոդել - 1PLM),
 - 2 պարամետրով միանյութ մոդել (2PLM),
 - 3 պարամետրով միանյութ մոդել (3PLM),
 - Մասնակի միավորներով բազմանյութ մոդել (PCM),
 - Ընդլայնարձակած մասնակի միավորներով բազմանյութ մոդել (GPCM):

At the bottom, there is a button labeled 'ՎԵՐԼՈՒԾԵԼ'.

Նկար 2. Համակարգի օգտվողի ինտերֆեյսի գլխավոր էջի տեսքը:

«Թեստի վերլուծություն» բաժնի միջոցով օգտվողը կարող է մուտքագրել (վերբեռնել) վերլուծվող թեստի փորձարկման արդյունքների և թեստի պատասխանների բանալիների ֆայլը՝ TXT կամ CSV ձևաչափով, ընտրել վերլուծության համար անհրաժեշտ համապատասխան մոդելները: «Որակական բնութագրեր» բաժնում օգտվողը կարող է ծանոթանալ թեստերի որակական բնութագրերի հակիրճ նկարագրությանը: Օգտվողին հնարավորություն է տրվում «Օրինակներ» բաժնից ներբեռնել փորձարկման համար անհրաժեշտ թեստերի արդյունքների և պատասխանների ֆայլերի ցուցադրական օրինակներ: «Օգնություն բաժինը»

նախատեսված է համակարգի օգտագործման նպատակով օգտվողին տեղեկատվություն տրամադրելու համար: Այստեղ բերված է թեստերի փորձարկումների արդյունքների և պատասխանների բանալիների ֆայլերի կառուցվածքի վերաբերյալ մանրամասն նկարագրություն:

Բաժին 4.4-ում նկարագարվել է թեստի որակի վերլուծության և գնահատման մոդուլը: Այս մոդուլը որպես մուտքային ինֆորմացիա ընդունում է ինտերֆեյսից օգտվողի կողմից վերբեռնված թեստի համապատասխան ֆայլերը, ինչպես նաև ընտրված վերլուծության պարամետրերի արժեքները: Այս մոդուլի միջոցով հաշվարկվում են մի շարք վիճակագրական ցուցանիշներ, Java psychometrics API գրադարանի հիման վրա հաշվարկվում են CTT և IRT ցուցանիշները, ինչպես նաև հաշվարկվում են ինֆորմացիայի տեսությունից ներմուծված ցուցանիշները: Կատարվում է սահմանային ընդունելի արժեքներ ունեցող ցուցանիշների գնահատում և առանձնացում, հաշվարկված ցուցանիշների հիման վրա տվյալների բազայում ստեղծվում են աղյուսակներ և գրանցվում հաշվարկված ցուցանիշների արժեքները:

Կարևոր է նշել, որ նկարագրված մոդուլը կարող է գործարկվել առանձնացված համակարգչի վրա որպես desktop ծրագիր: Այս տարբերակում օգտվողն անմիջապես առնչվում է մոդուլի GUI ինտերֆեյսի հետ, որը հնարավորություն է տալիս ինտերակտիվ ռեժիմում աշխատել մուտքագրված թեստավորման արդյունքների մատրիցայի հետ, կազմակերպել թեստի որակի վերլուծություն, գեներացնել և պահպանել տարբեր աղյուսակներ և գրաֆիկներ:

Բաժին 4.5-ում նկարագարվել է արդյունքների մշակման և հաշվետվության ձևավորման մոդուլը: Այս մոդուլի միջոցով ձևավորվում է հաշվետվությունների արխիվային ֆայլ-փաթեթը, որն իր մեջ ներառում է հաշվետվության rtf ֆայլը, գեներացված գրաֆիկները, աղյուսակների CSV ձևաչափով ֆայլերը: Այս փաթեթը օգտվողը կարողանում է ներբեռնել օգտվողի ինտերֆեյսից տրամադրվող հղման միջոցով: Հաշվետվության ֆայլի rtf ձևաչափն ունի մի շարք առավելություններ, որոնցից կարելի է նշել մի քանիսը.

- հեշտ կիրառելի է, կարելի այն բացել տարբեր տեքստային խմբագրիչներով,
- առանց որևէ լրացուցիչ փոփոխությունների կարելի է հեշտությամբ տպել,
- հեշտությամբ կարելի է խմբագրել ֆայլի պարունակությունը:

Հաշվետվության rtf ձևաչափով ֆայլը բաղկացած է հետևյալ բաժիններից.

- ներածություն,
- ընդհանրացված վիճակագրություն,
- CTT ցուցանիշներ,
- IRT ցուցանիշներ,

- ինֆորմացիայի տեսության ցուցանիշներ,
- գրաֆիկներ:

Ներածություն բաժնում բերված է ընդհանուր տեղեկություններ հաշվետվության բովանդակության, ինչպես նաև թեստերի որակական ցուցանիշների վերաբերյալ: Յուրաքանչյուր բաժնի սկզբում բերված է այդ բաժնում ներկայացվող աղյուսակների և տվյալների վերաբերյալ ընդհանուր տեղեկատվություն: Այն ցուցանիշները, որոնք ունեն սահմանային արժեքներ, դրանց համապատասխան տրվում է գունավորում (կարմիր գույն), ինչը թույլ է տալիս օգտվողին միանգամից տարբերակել խնդրահարույց առաջադրանքները կամ թեստի այլ ցուցանիշներ: Առաջադրանքների էնտրոպիայի, պայմանական էնտրոպիայի և միջին փոխադարձ ինֆորմացիայի մեծությունների արժեքների աղյուսակները գեներացվում են առանձին CSV ֆայլի տեսքով:

Բաժին 4.6-ում նկարագրված է համակարգի մշակման տեխնիկական լուծումները: Համակարգը մշակված է կլիենտ-սերվեր ճարտարապետությամբ, որը հնարավորություն է տալիս ապահովել տարբեր տեխնոլոգիաներով մշակված բաղադրիչների փոխգործակցությունը, ինչպես նաև օգտվողից քողարկել մաթեմատիկական և վիճակագրական հաշվարկներն իրականացնող մոդուլը: Համակարգը կարելի է տեղադրել PHP և JRE սատարումով Web սերվերի վրա: Համակարգի առաջին բաղադրիչը web հենքով օգտվողի ինտերֆեյսն է: Նման լուծումը բավարարում է համակարգի պլատֆորմից անկախ լինելու և լոկալ ցանցի կամ համացանցի միջոցով հասանելի լինելու պահանջը: Web օգտվողի ինտերֆեյսի մշակման համար օգտագործվել են հետևյալ տեխնոլոգիաները՝ HTML, CSS, JavaScript: Սերվերային մասը մշակվել է PHP ծրագրավորման լեզվով: PHP սերվերային սկրիպտը հանդիսանում է կապող կամուրջ օգտվողի web ինտերֆեյսի և համակարգի երրորդ ամենածավալուն բաղադրիչ հանդիսացող՝ մաթեմատիկական և վիճակագրական հաշվարկներն իրականացնող Java հավելվածի միջև: Մաթեմատիկական և վիճակագրական հաշվարկներն իրականացնող Java հավելվածում օգտագործվում են մի շարք գրադարաններ, որոնցից կարևոր է նշել բաց ծրագրային կոդով Java psychometrics API գրադարանը: Այս գրադարանը պարունակում է IRT և CTT տեսությամբ հաշվարկներ կատարելու համար անհրաժեշտ դասեր: Java հավելվածը մուտքային ֆայլերի տվյալները պահպանում է Java Apache Derby բաց ծրագրային կոդով ռեյլաբազային տվյալների բազայում, ինչն էլ թույլ է տալիս չսահմանափակել մուտքագրվող փորձարկվողների կամ առաջադրանքների քանակը: Տվյալների բազայում պահպանվում են նաև հաշվարկվող ցուցանիշների աղյուսակները: Վերջնական հաշվետվությունների ֆայլը գեներացվում և ձևավորվում է արդյունքների մշակման և հաշվետվության ձևավորման մոդուլի միջոցով: Այս մոդուլում գրաֆիկների գեներացման համար ներառվել է JFreeChart անվճար գրադարանը, որը տրամադրում է

ճկուն միջոցներ գրաֆիկներ ստեղծելու համար: Այն սատարում է տարատեսակ երկչափ և եռաչափ գրաֆիկներ, գրաֆիկների ֆայլերի JPG, PNG, SVG և այլ ձևաչափեր, ունի շատ մանրամասն ու մատչելի նկարագրություններ: Հաշվետվությունների ֆայլը տրվում է rtf ձևաչափով, որի ձևավորման համար օգտագործվում է Apache POI API գրադարանը, կիրառվում են այս գրադարանի XWPF (XML Word Processor Format) բաղադրիչի մեթոդները :

Բաժին 4.7-ում ամփոփված են համակարգի հիմնական և իրականացված տեխնիկական լուծումների առանձնահատկությունները: Դրանք են՝

- մշակված համակարգն իրականացնում է թեստերի որակի գնահատում միաժամանակ CTT-ի, IRT-ի տեսությունների և մեր կողմից առաջարկված նոր մոտեցման հիման վրա,
- օգտվողի ինտերֆեյսը web հենքով է, ինչը թույլ է տալիս օգտվողին դիմել համակարգին լոկալ ցանցի միջոցով կամ համացանցի միջոցով՝ գործարկելով այն web զննարկիչի միջավայրում,
- օգտվողի ինտերֆեյսը բոլոր փաստաթղթերը հասանելի են հայերեն լեզվով,
- գեներացվող հաշվետվության ֆայլը ձևավորված rtf փաստաթուղթ է, որն իր մեջ ներառում է ձևավորված աղյուսակներ, գրաֆիկներ, ինչպես նաև հաշվարկված և գնահատված ցուցանիշների վերաբերյալ նկարագրություններ և ցուցումներ,
- որպես այլընտրանք ստեղծված է հնարավորություն տրամադրել աղյուսակները CSV ձևաչափով, իսկ գրաֆիկները JPG ձևաչափով,
- պարզեցված օգտվողի ինտերֆեյսը հնարավոր է դարձնում համակարգի կիրառումը օգտվողների լայն շրջանակների համար,
- համակարգը պլատֆորմից անկախ է, ինչի արդյունքում էլ այն կարող է գործարկվել տարբեր օպերացիոն համակարգերի միջավայրերում,
- քանի որ Java հավելվածները պլատֆորմից անկախ են, համակարգի բաղադրիչ հանդիսացող վիճակագրական և մաթեմատիկական ցուցանիշները հաշվարկող մոդուլը կարող է կիրառվել առանձին, որպես desktop հավելված:

Առենախոսության **հինգերորդ՝ «Նոր մեթոդի բնութագրիչների ընդունելի արժեքների որոշում»** վերնագրով զլուխը նվիրված է առաջարկված թեստերի որակի գնահատման նոր մոդելի փորձարկումների նկարագրությանը և արդյունքների վերլուծությանը: Փորձարկումների նպատակն էր փորձնական ճանապարհով որոշել հաշվարկվող ցուցանիշների սահմանային թույլատրելի արժեքները, ինչպես նաև վավեր վերլուծության իրականացման համար փորձարկվողների բավարար քանակը:

Բաժին 5.1-ում ներկայացված է կատարված փորձարկումների նկարագրությունը: Փորձարկման համար ընտրվել են 2008 թվականի «Հայոց լեզու և գրականություն» առարկայի ավարտական միասնական քննությունների թեստավորման արդյունքները, որոնք ծածկագրված տեսքով տրամադրվել է ՀՀ կառավարության աշխատակազմի «Գնահատման և թեստավորման կենտրոն» ՊՈԱԿ-ի կողմից: Վերլուծության համար ընտրվել է 4 թեստ, որոնցից յուրաքանչյուրը պարունակում է 80 առաջադրանք, իսկ թեստավորմանը մասնակցած փորձարկվողների ընդհանուր թիվը 2000 է: Վերլուծության համար օգտագործվել է մեր կողմից մշակված թեստերի որակի գնահատման ծրագրային համակարգը: Չորս թեստերի յուրաքանչյուր առաջադրանքի համար հաշվարկվել են $H(X_i)$, $I(X_i \wedge X_j)$, CTT բարդության ինդեքսի, $R(X_i, X_j)$ կոռելյացիոն գործակցի և IRT b պարամետրի արժեքները: Ուսումնասիրվել է նշված պարամետրերի համադրելիությունը և որոշվել է $H(X_i)$ և $I(X_i \wedge X_j)$ պարամետրերի թույլատրելի սահմանային արժեքները: $H(X_i)$ և $I(X_i \wedge X_j)$ մեծությունների արժեքների կախվածությունը փորձարկվողների խմբի քանակից ուսումնասիրելու և համապատասխան փորձարկվողների բավարար քանակը որոշելու համար յուրաքանչյուր թեստի համար կատարվել է հինգ փորձարկում: Յուրաքանչյուր փորձարկման համար պատահական ընտրությամբ առանձնացվել է համապատասխանաբար 500, 300, 200, 100 և 50 փորձարկվողի արդյունք:

Բաժին 5.2-ում կատարված են եզրակացություններ առաջարկված մոդելի փորձարկումների վերաբերյալ: Մեր կողմից առաջարկված մեթոդները ճշգրիտ գնահատում են թեստերի որակը և համադրելի են CTT և IRT գնահատման մեթոդների հետ: IRT տեսության համեմատությամբ առավել պարզ մաթեմատիկական վերլուծություն է կատարվում: Առաջադրանքերի միջին փոխադարձ ինֆորմացիան ունի համարժեքը CTT-ում, իսկ IRT-ում համարժեքը բացակայում է: IRT-ի և CTT-ի համեմատությամբ ավելի փոքր քանակությամբ թեստավորվողների խումբ է հարկավոր վավեր վերլուծություն և գնահատում իրականացնելու համար:

Աշխատանքի հիմնական արդյունքները:

Այսպիսով, աշխատանքում ստացվել են հետևյալ հիմնական արդյունքները.

- Մշակվել է թեստերի որակի գնահատման նոր մեթոդ՝ հիմնված ինֆորմացիայի տեսության տարրերի վրա [4]:
- Որոշվել են առաջարկված մեթոդի ցուցանիշների ընդունելի սահմանները և վավեր վերլուծության համար անհրաժեշտ թեստավորման արդյունքների քանակը, որը էականորեն ավելի փոքր է մյուս մեթոդների հետ համեմատած [5][1]:
- Մշակվել է CTT, IRT տեսությունների և առաջարկված նոր մեթոդի հիման

վրա թեստերի որակի վերլուծություն և գնահատում իրականացնող «Թեստերի որակի վերլուծություն» ծրագրային համակարգը, որը գործարկվում է սերվերի վրա, ունի web հենքով օգտվողի պարզ և մատչելի ինտերֆեյս, տրամադրում է rtf ձևաչափի վերլուծության հաշվետվության ֆայլ [3] [2]:

- Մշակվել է վիճակագրական և մաթեմատիկական ցուցանիշները հաշվարկող մոդուլի GUI ինտերֆեյս, ինչի շնորհիվ էլ այն կարող է կիրառվել նաև առանձին համակարգչի վրա որպես desktop ծրագրային հավելված: Մշակված մոդուլը բազմապլատֆորմ է և գործարկվում է տարբեր օպերացիոն համակարգերի միջավայրերում [3]:

Հրատարակված աշխատությունների ցանկը

1. Ավետիսյան Վ., “Գիտելիքների ստուգման թեստերի որակական բնութագրերի հետազոտություն”, Հայաստանի Ճարտարագիտական Ակադեմիայի Լրաբեր (ՀՃԱԼ), Երևան, Հ.12, N 1, 2015, էջ 156-163:
2. Avetisyan V., “Survey of software for the test quality analysis”, International Journal Information Content and Processing, Bulgaria, ITHEA, Volume 2, Number1, 2015, pp. 82-92.
3. Haroutunian M., Avetisyan V., “Development of the test quality evaluation system”, Proceedings of the International Conference on Computer Science and Information Technologies (CSIT 2015), Armenia, Yerevan, 2015, September 28-October 2, pp. 372-375.
4. Haroutunian M., Avetisyan V., “New approach for test quality evaluation based on Shannon Information measures”, Transactions of IIAP of the NAS RA, Mathematical Problems of Computer Science, Yerevan, vol. 44, 2015, pp. 7-21.
5. Haroutunian M., Avetisyan V., “Analysis of the Experiments of a New Approach for Test Quality Evaluation”, Transactions of IIAP of the NAS RA, Mathematical Problems of Computer Science, Yerevan, vol. 45, 2016, pp. 35-43.

РАЗРАБОТКА МЕТОДА И ПРОГРАММНОЙ СИСТЕМЫ ОЦЕНКИ КАЧЕСТВА ТЕСТОВ ДЛЯ ПРОВЕРКИ ЗНАНИЙ

РЕЗЮМЕ

Актуальность исследования: Мировой опыт свидетельствует об эффективности экзаменационных тестов, используемых для проверки знаний. В последние годы в Республике Армения тоже широко применяются технологии тестирования. Важно отметить, что объективность тестирования и оценка знаний в значительной степени зависят от качества тестов.

Для оценки качества тестов статистическим способом, существуют теории, среди которых наиболее широко используется классическая теория тестов (Classical Test Theory-CTT) и современная математическая теория тестов (Item Resonse Theory-IRT). Стоит отметить, что статистический аппарат достаточно сложен для преподавателей различных областей, а расчеты трудоемки для узкого специалиста. Для решения этой задачи, разрабатываются системы, программные пакеты, посредством которых осуществляется анализ результатов экспериментального тестирования, и оценка качественных характеристик теста.

Программные пакеты, в основном, не имеют простой интерфейс, ориентированный на пользователя, не предоставляют отчет анализа результатов, выполнены на английском языке и имеют высокую цену.

В этой ситуации актуально была разработка более простого метода оценки качества тестов и программного обеспечения на армянском языке.

Цель исследования: Для оценки качества тестов была поставлена задача – разработать новый метод и программное обеспечение, которое обеспечит всесторонний анализ с поддержкой армянского языка.

Научная новизна исследования

1. Разработан новый подход для оценки качества тестов, основанный на элементах теории информации, который обеспечивает более быстрые расчеты и для осуществления правомерного анализа требует гораздо меньше результатов тестирования.
2. Разработанное программное обеспечение осуществляющее анализ качества и оценки тестов доступно на армянском языке, работает на сервере и имеет веб-интерфейс пользователя, что и позволяет применить систему посредством локальной сети или интернета, запуская его в среде веб-браузера.
3. Модуль расчетов статистических и математических показателей являющийся отдельным компонентом системы, может быть применен также в качестве desktop приложения.
4. Разработанная система не зависит от типа операционной системы сервера.

Практическая значимость исследования: Разработанная система с простым и доступным веб-интерфейсом, позволяет пользователю без владения сложным аппаратом статистической теории тестов и без установки дополнительных программных средств, организовать анализ качества и оценки разработанных тестов.

С помощью системы генерируется отчет анализа качества и оценки тестов, который является файлом формата rtf (rich text format) и содержит как исчисляемые и оцениваемые показатели, таблицы, графики, так и описания, и указания по ним.

Система может быть применена педагогами различных специальностей, организациями, занимающимися проведением различных тестирований (учебных заведений, центров тестирования и т. д.), с целью анализа результатов тестов и оценки их качества.

Результаты: В работе получены следующие основные результаты:

- Разработан новый метод оценки качества тестов, основанный на элементах теории информации [4].
- Были определены приемлемые границы показателей предложенного метода [5].
- Оценено количество необходимых результатов тестирования для правомерного анализа, которое существенно меньше по сравнению с другими методами [1, 5].
- Разработано программное обеспечение «Анализ качества тестов» на двух языках, осуществляющее анализ качества и оценки тестов, на основе теорий СТТ, IRT и нового метода, которое работает на сервере, имеет простой и удобный веб интерфейс, предоставляет файл анализа отчетности формата rtf [2, 3].
- Разработан GUI-интерфейс для модуля расчетов статистических и математических показателей, благодаря чему он может быть применен также для отдельных компьютеров в качестве desktop приложения. Разработанный модуль кроссплатформенный и работает в среде различных операционных систем [3].

DEVELOPMENT OF KNOWLEDGE TESTS QUALITY EVALUATION METHOD AND SOFTWARE SYSTEM

SUMMARY

The relevance of the research

International experience makes obvious the effectiveness of examination computer tests applied to test knowledge. In recent years a wide range of testing technologies are applied in Armenia as well. It is important to mention that the objectivity of computer testing and knowledge assessment greatly depends on the quality of tests.

To evaluate the test quality in a statistical method some existing theories can be used, among which Classical Test Theory (CTT) and modern mathematical test theory (Item Resonse Theory-IRT) are widely used.

It should be noted that the statistical apparatus is quite difficult for pedagogues in various fields, and the calculations require much time from a specialist in a certain field. To solve this problem the systems, software packages are developed through which an experimental test results' analysis and qualitative characteristics evaluation of the test are implemented.

The software packages do not have a user friendly simple interface, they do not provide the report on analysis results, they are in the English language and are expensive.

In this situation it is important to develop a simple method and a software system supported Armenian language for test quality evaluation.

The aim of the research

To organize the checking process of tests quality as well as the assessment process a new method and software system should have been developed to provide a comprehensive analysis with the Armenian language support.

Scientific novelty of the research

1. Based on the elements of information theory, a new approach to assess the quality of tests has been developed which provides faster calculations and requires much number of testing results to implement a valid analysis.
2. The software system, with the support of the Armenian language and directed to evaluating the qulaity of the developed tests, is run on the server and has a web-based user interface which allows to apply to the system through the local network or via the Internet launching it in the web browser environment.
3. The module which calculates the statistical and mathematical indicators that are considered to be a separate component of the system can also be used as a computer desktop application.
4. The developed system is not dependent on the type of operating system of the server.

The practical significance of the research

The developed system having the Armenian language support, with a simple and

accessible user web interface enables the users to conduct the analysis of the developed tests quality as well as evaluation without mastering the complicated apparatus of statistical theory of the tests and without installing additional software resources.

The report on analysis and evaluation of tests quality is generated through the system. The report is an editable file of rich text format (rtf) and contains not only figures, tables, graphs that can be calculated and estimated but also descriptions and instructions on them.

The system can be used by various professionals and pedagogues engaged in test development process as well as by different organizations (educational institutions and test centers, etc.) engaged in the testing process aimed at test results' analysis and tests quality evaluation.

The results of the research

Thus, the main results of the work are the following:

- A new method aimed at the tests quality evaluation based on the elements of information theory has been developed [4].
- The acceptable limits to the suggested method's indicators and the amount of test results required for a valid analysis are identified. The above mentioned amount is significantly less compared to other methods [1, 5].
- The software system of "Test Quality Analysis" has been developed. The system implements test quality analysis and evaluation based on the suggested new method and CTT, IRT theories. The software system is run on the server, it has a web-based simple and accessible user interface, provides the analysis report in rich text format (rtf) file [2, 3].
- The GUI interface of calculating module of the mathematical and statistical indicators has been developed. It can be applied on a computer desktop as a desktop software application. The developed module is cross-platform and is applied in various operating system environments [3].

