ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

ՆԱՐԵԿ ՍԱՄՍՈՆԻ ՓԱՀԼԵՎԱՆՅԱՆ

ԳԱՂՏՆԻՔ ԳԵՆԵՐԱՑՆՈՂ ԿԵՆՍԱՉԱՓԱԿԱՆ ՀԱՄԱԿԱՐԳԻ ԻՆՖՈՐՄԱՑԻՈՆ-ՏԵՍԱԿԱՆ ՀԵՏԱԶՈՏՈՒԹՅՈՒՆ ԵՎ ՆՈՐ ՓԱԹԵԹԻ ՄՇԱԿՈՒՄ R ՄԻՋԱՎԱՅՐԻ ՀԱՄԱՐ

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

НАРЕК САМСОНОВИЧ ПАЙЛЕВАНЯН

ИНФОРМАЦИОННО ТЕОРЕТИЧЕСКИЙ АНАЛИЗ БИОМЕТРИЧЕСКОЙ СИСТЕМЫ С ГЕНЕРИРУЕМЫМ СЕКРЕТОМ И СОЗДАНИЕ НОВОГО МОДУЛЯ ДЛЯ СРЕДЫ R

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 տեխ.գիտ.թեկնածու	Հ.Բ.Մարանջյան Մ.Ղ.Գյուրջյան
Առաջատար կազմակերպություն՝	Հայաստանի ազգային պոլիտեխնիկական համալսարան	

Պաշտպանությունը կայանալու է 2016թ. հունիսի 10-ին ժամը 17:00 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. Սարուխանյան

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
	кандидат тех.наук	М.К.Гюрджян
Ведущая организация:	Национальный Политехнический Университет Армении	

Защита состоится 10-го июня 2016г. в 17:00 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 ознакомится 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

Վերջին տարիներին անվտանգության միջոցների մշակման կարևոր բաժին է կազմում կենսաչափական տվյալների վրա հիմնված բազմազան համակարգերի հետազոտումը: Կենսաչափական անվտանգության համակարգերը հիմնվում են մարդու ֆիզիկական կամ վարքագծային բնութագրիչների վրա, ինչպիսիք են դեմքը, մատնահետքերը, ձայնը, աչքի ցանցաթաղանթը, արյունատար անոթները, քայլվածքը և այլն[1]: Կենսաչափական տվյալների առավելությունը ավանդական ծածկագրման միջոցներից կայանում է նրանում, որ դրանք եզակի նույնացուցիչներ են անձի համար[2], որոնք չի կարելի կորցնել կամ կեղծել: Սակայն միակությունը իր հերթին նոր խնդիրներ է առաջացնում, քանի որ անհրաժեշտության դեպքում կենսաչափական տվյալները հնարավոր չէ փոխել: Այս պրոբլեմները ստիպում են դիտարկել ավելի բարդ համակարգերը:

Ցանկացած կենսաչափական համակարգ ինֆորմացիոն-տեսական տեսանկյունից կարող է ապահովել միայն գաղտնիության որոշակի մակարդակ[3]: Կիրառություններից կախված դիտարկվում են կենսաչափական նույնականացման և գաղտնիք գեներացնող տարբեր մոդելներ: Յուրաքանչյուր մոդելի համար առաջին կարևոր խնդիրն է որոշել նույնականացման կամ գաղտնիքի հասանելի արագությունը: Ինֆորմացիայի տեսության երկրորդ խնդիրը հուսալիության ֆունկցիայի հետազոտումն է: Այն իրենից ներկայացնում է հիմնական բնութագրիչների (գաղտնիքի արագության, հուսալիության, սխալի հավանականության, արտահոսքի արագության) փոխկապվածությունը արտահայտող ֆունկցիան և ընդհանրացնում է լավագույն հասանելի արագության գաղափարը: Այս ֆունկցիայի ուսումնասիրությունը բավականին բարդ խնդիր է, դրա մասին է վկայում այն փաստը, որ նույնիսկ պարզագույն ընդհատ առանց հիշողության կապուղու համար ֆունկցիայի ճշգրիտ տեսքը հայտնի չէ, հաջողվում է գտնել միայն վերին և ստորին գնահատականները[4]:

Բացի տեսական հետազոտությունների բարդությունից առաջ են գալիս նաև ստացված տեսական արդյունքների կիրառության դժվարություններ, քանի որ ստացված մաթեմատիկական բանաձևերը դժվար հաշվարկելի են: Այսպիսով,

---

[1] A. Jain, R. Bolle, and S. Pankanti. "*Biometrics: Personal Identification in a Networked Society",* Kluwer Academic Publishers, 1999.

[2] R. Clarke, "*Human identification in information systems: Management challenges and public policy issues.*", Information Technology & People, T. 7, pp. 6-37, 1994.

[3] T. Ignatenko and F. Willems, "Biometric security from an information-theoretical perspective," Foundations and Trends in Communications and Information Theory, vol. 7, no. 2-3, 2012.

[4] E. Haroutunian, M. Haroutunian and A. N. Harutyunyan, "Reliability criteria in information theory and in statistical hypothesis testing," Foundations and Trends in Communications and Information Theory, vol. 4, no. 2-3, 2008.

նմանատիպ հաշվարկներ իրականացնող ծրագրային միջոցի մշակումը արդիական խնդիր է:

Վերջին տարիներին արագ զարգանում է R ծրագրավորման միջավայրը, ինչը պայմանավորված է մի շարք առավելություններով[5]: R-ը տվյալների հավաքագրման, վիճակագրական հետազոտությունների, ժամանակային շարքերի վերլուծության, մաթեմատիկական հաշվարկներ իրականացնելու և համապատասխան գծապատկերներ կառուցելու միջավայր է: R-ը օգտագործողին տրամադրում է տարբեր հնարավորություններ վիճակագրական գծային և ոչ գծային մոդելավորման, ժամանակային շարքերի վերլուծության, տվյալների դասակարգման, կլաստերիզացիայի համար: Այն կարող է նպաստել նաև ինֆորմացիայի տեսության բարդ բանաձևերի հաշվարկմանը, քանի որ այդ բանաձևերում պահանջվում է դիտարկել մեծ թվով բաշխումներ: Սակայն R-ում գոյություն ունեցող ինֆորմացիայի տեսության փաթեթը իրականացնում էր սահմանափակ գործառույթներ: Ուստի խնդիր դրվեց մշակել ավելի կատարելագործված նոր փաթեթ ինֆորմացիայի տեսության խնդիրների լուծման համար:

Հետազոտության նպատակը

Աշխատանքի նպատակն է ինֆորմացիայի տեսության ֆունկցիաների բարդ հաշվարկների իրականացման համար մշակել նոր փաթեթ R միջավայրի համար: Այն իրականացնելու համար խնդիր դրվեց ինֆորմացիայի-տեսության տեսանկյունից հետազոտել որպես կիրառական օրինակ՝ գաղտնիք գեներացնող կենսաչափական մոդելի հիմնական բնութագրիչների (գաղտնիքի արագության, սխալի հավանականության ցուցչի կամ հուսալիության, արտահոսքի արագության) փոխկապվածությունն արտահայտող ֆունկցիաները:

Հետազոտման օբյեկտը

Ուսումնասիրության օբյեկտներն են կենսաչափական նույնականացման մոդելի *E*-հասանելի գաղտնիքի արագության ֆունկցիան, որը հասանելի գաղտնիքի արագության ընդհանրացումն է, ինչպես նաև R միջավայրում նոր փաթեթների ստեղծման սկզբունքները, C++ ծրագրավորման լեզվով R-ում նոր փաթեթների ինտեգրումը և բազմահոսքային ալգորիթմների աշխատանքի ապահովման հնարավորությունները R միջավայրում:

Հետազոտման մեթոդները

Ատենախոսության մեջ կիրառվում է ինֆորմացիայի տեսությունը: Օգտագործվում են R-ում C++-ի գրադարանների ինտեգրման մեթոդները, NVidia CUDA տեխնոլոգիայի[6] գրադարանի կիրառմամբ գրաֆիկական քարտի հետ աշխատանքի մեթոդներ, ինչպես նաև կլաստերում գործարկման համար OpenMPI բաց ծրագրային

[5] D. Smith, "The R ecosystem," in The R User Conference (useR!), 2011.

[6] Nvidia CUDA Technology, http://www.nvidia.com/CUDA

կոդով գրադարանի[7] բարդ տվյալների տիպերի կառուցման մեթոդներ:

Հետազոտության գիտական նորույթը

1. Գաղտնիք գեներացնող կենսաչափական մոդելի համար կառուցվել են *E*-հասանելի գաղտնիքի արագության ստորին և վերին գնահատականները, որոնք, երբ $E \to 0$, համընկնում են և հավասարվում են հասանելի գաղտնիքի արագության հետ:
2. Ներկայացվել է փոփոխելի հուսալիության աստիճան ունեցող գաղտնիք գեներացնող կենսաչափական համակարգի մոդել:
3. Մշակվել է ինֆորմացիայի տեսության բարդ ֆունկցիաների հաշվարկներ իրականացնող նոր մոդուլ R միջավայրի համար:
4. Մշակված մոդուլը ապահովում է հաշվարկի երեք տեսակի զուգահեռացման հնարավորություն և թույլ է տալիս փոփոխել զուգահեռացման տեսակը հաշվարկի ընթացքում:
5. Փաթեթում բազմահոսքայնության հետ առնչվող խնդիրների լուծման համար առաջարկվել են փականքների (locks) և փակուղիներից (deadlocks) խուսափման մեխանիզմներ:

Ստացված արդյունքների կիրառական նշանակությունը

Մշակված մոդուլը հնարավորություն է տալիս հաշվելու ոչ միայն էնտրոպիա, տարամիտություն, միջին փոխադարձ ինֆորմացիա, այլ նաև այդ ֆունկցիաների փոքրագույն արժեքները, ըստ որոշակի պայմանների:

Մոդուլը առաջարկում է զուգահեռացման 3 հնարավորություն՝ զուգահեռացում տեղային պրոցեսորի միջոցով (CPU տիպ), զուգահեռացում գրաֆիկական քարտի միջոցով (GPU տիպ) և զուգահեռացում կլաստերի միջոցով (MPI տիպ):

Մոդուլը կարող է կիրառվել ինֆորմացիայի տեսության մասնագետների կողմից մի շարք բարդ ֆունկցիաներ հաշվարկելու համար:

Ներդրումներ

Մշակված փաթեթը անցել է փորձաքննություն «CRAN team» և «R Foundation» հանձնաժողովների կողմից և ներդրվել է R միջավայրի «The Comprehensive R Archive Network» փաթեթների արխիվի պաշտոնական ցանցում:

Պաշտպանության ներկայացվող հիմնական դրույթները

- Գաղտնիք գեներացնող կենսաչափական մոդելի համար կառուցվել են *E*-հասանելի գաղտնիքի արագության ստորին և վերին

[7] Gabriel E., Fagg G., Bosilca G., "Open MPI: Goals, concept, and design of a next generation MPI implementation," in Proceedings of 11th European PVM/MPI Users' Group Meeting, Budapest, Hungary : pp 97-104, 2004.

գնահատականները, որոնք փոքր E-երի դեպքում համընկնում են և հավասարվում են հայտնի հասանելի գաղտնիքի արագության հետ:

- Ներկայացվել է փոփոխելի հուսալիության աստիճան ունեցող գաղտնիք գեներացնող կենսաչափական համակարգի մոդել:
- Մշակվել է նոր փաթեթ R միջավայրի համար, որը իրականացնում է Ինֆորմացիայի տեսության բարդ ֆունկցիաների հաշվարկներ:
- Փաթեթում տրվել են հաշվարկների իրականացման երեք տեսակի զուգահեռացման հնարավորություններ՝ զուգահեռացում տեղային պրոցեսորի միջոցով (CPU տիպ), զուգահեռացում գրաֆիկական քարտի միջոցով (GPU տիպ) և զուգահեռացում կլաստերի միջոցով (MPI տիպ):
- Փաթեթում բազմահոսքայնության հետ առնչվող խնդիրների լուծման համար առաջարկվել են փականքների (locks) և փակուղիներից (deadlocks) խուսափման մեխանիզմներ:

Ապրոբացիա

Ատենախոսության արդյունքները զեկուցվել են՝

- Հայկական մաթեմատիկական միության նստաշրջանում, Երևան, 2014 թ.
- ԳՊՄԻ 80 ամյակին նվիրված հանրապետական գիտաժողովում, Գյումրի, 2014 թ.
- На международной научно-технической конференции студентов и молодых специалистов из стран-участниц РСС по направлениям «Информационные технологии и системы связи», Москва, 2015г.
- 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 թ.,
- ինչպես նաև ՀՀ ԳԱԱ ԻԱՊԻ ընդհանուր սեմինարում:

Հրապարակումներ

Ատենախոսության հիմնական արդյունքները հրապարակված են 6 գիտական աշխատություններում, որոնք թվարկված են սեղմագրի վերջում [1-6]:

Աշխատանքի կառուցվածքը և ծավալը

Ատենախոսությունն իր մեջ ներառում է ներածություն, 5 գլուխներ՝ իրենց ամփոփումներով, եզրակացություն, գրականության ցանկ՝ իր 104 հղումներով: Ատենախոսության ընդհանուր ծավալը 101 էջ՝ 1 աղյուսակով և 21 պատկերներով:

ԱՇԽԱՏԱՆՔԻ ՀԻՄՆԱԿԱՆ ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆԸ

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

Ատենախոսության ***առաջին*** գլխում նկարագրվել են ինֆորմացիայի տեսության գաղափարները, աշխատանքի հիմնական խնդիրները, ձևակերպումները: Ներկայացվել են մուտքի վերահսկման կենսաչափական համակարգերը, այդ համակարգերի գաղտնիության տեսակները և առկա խնդիրները:

Բաժին 1.1-ում դիտարկվել է պարզագույն կապի համակարգի Շենոնի մոդելը[8] և այդ մոդելի մաթեմատիկական նկարագրման համար կատարվել են որոշակի նշանակումներ: Տրվել են Շենոնի կողմից ձևակերպված օպտիմալ հաղորդման երկու հիմնախնդիրները, որոնք են՝ տալ աղբյուրի ստեղծած ինֆորմացիայի չափման եղանակ և պարզել, թե աղբյուրից եկող ինֆորմացիայի ո՞ր քանակը կարելի է հաղորդել կապուղով՝ ապահովելով հաղորդման բարձր որակը (սխալի փոքր հավանականությունը):

Բաժիններ 1.2 - 1.4-ում ձևակերպվել են ինֆորմացիայի տեսության հիմնական գաղափարները և դրանց հատկությունները մասնավորապես՝ միջին փոխադարձ ինֆորմացիայի, էնտրոպիայի, Կուլբակի-Լեյբլերի տարամիտության:

Բաժին 1.5-ում ձևակերպվել է Շենոնի հիմնական հայտնագործություններից մեկը՝ Շենոնի կոդավորման թեորեմը կապուղու համար: Ըստ որի ամեն մի կապուղի կարելի է բնութագրել մի թվով, որը կոչվում է *ունակություն* և նշանակվում է C-ով: Շենոնի թեորեմը ապացուցված է բավականաչափ լայն դասերի կապուղիների համար, այդ թվում՝ ընդհատ առանց հիշողության կապուղու համար: Այդ կապուղու դեպքում, ինչպես ցույց է տվել Շենոնը, ունակությունը արտահայտվում է փոխադարձ ինֆորմացիայի միջոցով, ընդ որում ունակությունը նույնն է սխալի առավելագույն և միջին հավանականության դեպքերում:

Բաժին 1.6-ում նկարագրվել է ինֆորմացիայի տեսության արդյունքների ստացման կարևորագույն եղանակներից մեկը՝ կազմերի մեթոդը: Դիտարկվել են համատեղ կազմի, պայմանական կազմի գաղափարները, ձևակերպվել են կազմերի հիմնական հատկությունները:

Բաժին 1.7-ում ձևակերպվել են հուսալիության ֆունկցիայի, E-ունակության գաղափարները: Նկարագրվել են E-ունակության վերին և ստորին գնահատականները: Այդ ֆունկցիայի հետազոտումը բավականին բարդ խնդիր է: Դրա մասին է վկայում այն փաստը, որ նույնիսկ պարզագույն ընդհատ առանց հիշողության կապուղու դեպքում ֆունկցիայի ճշգրիտ տեսքը հայտնի չէ: Հայտնի են միայն ֆունկցիայի վերին և ստորին

[8] Shannon C. E, A mathematical theory of communication, The Bell System Technical Journal, Vol. 27, pp. 623-656, 1948.

գնահատականները[9], որոնք համընկնում են ունակությանը մոտ արագությունների դեպքում:

Բաժին 1.8-ում նկարագրվել են կենսաչափական մուտքի վերահսկման համակարգերը: Ձևակերպվել են ավանդական ծածկագրման համակարգերում առկա թերությունները: Ներկայացվել են կենսաչափական տվյալներից բխող սահմանափակումները:

Բաժին 1.9-ում նկարագրվել են ներկայումս գործածության մեջ գտնվող ավանդական կենսաչափական համակարգերը և նրանցում առկա սահմանափակումները, ձևակերպվել են կենսաչափական համակարգի հիմնական բնութագրիչները՝ կեղծ մերժման գործակիցը (ԿՄԳ) և կեղծ ընդունման գործակիցը (ԿԸԳ):

Բաժին 1.10-ում ձևակերպվել են գաղտնիության տեսակները: Նկարագրվել են գաղտնագրական հաղորդակարգերի անվտանգության գնահատման համար օգտագործվող հասկացությունները՝ ինֆորմացիոն-տեսական գաղտնիությունը և հաշվողական գաղտնիությունը: Ձևակերպվել է կատարյալ գաղտնիություն ապահովող համակարգի գաղափարը:

Բաժին 1.11-ում նկարագրվել են կենսաչափական գաղտնիություն ապահովող համակարգերը, մասնավորապես, դիտարկվել են կենսաչափական նույնականացման համակարգը և այդ համակարգը բնութագրող գաղափարները:

Ատենախոսության ***<u>երկրորդ</u>*** գլուխը նվիրված է R ծրագրավորման միջավայրի առանձնահատկությունների և հնարավորությունների նկարագրությանը: Այս գլխում հիմնավորված է R միջավայրի ընտրությունը ատենախոսության նպատակի իրականացման համար:

Բաժին 2.1-ում շարադրվել են R միջավայրի կիրառությունները ու հնարավորությունները, R-ի միջուկի ծրագրային կոդի վերլուծությունը:

Բաժին 2.2-ում ձևակերպվել են R միջավայրի և այլ վիճակագրական փաթեթների համեմատական հետազոտության արդյունքները: Բերված են R-ի առավելությունները այլ փաթեթների (SAS, STATA, SPSS, Matlab) նկատմամբ:

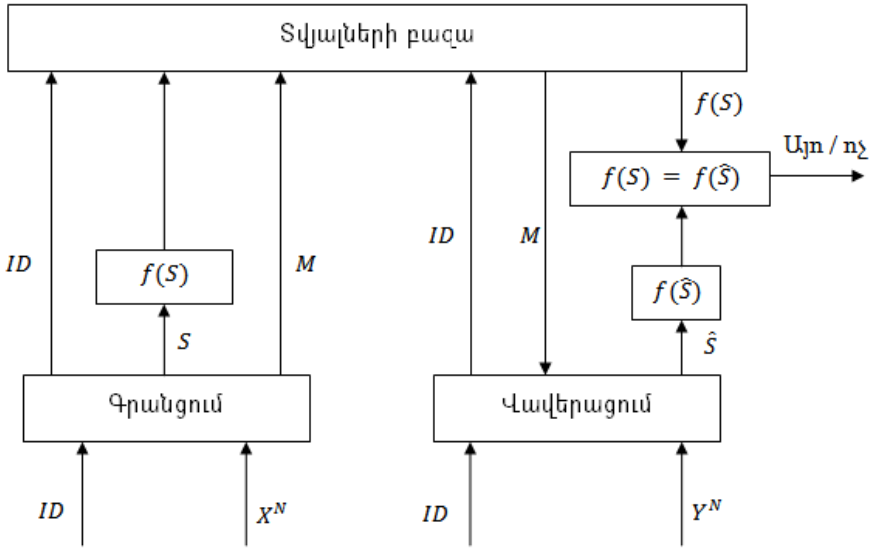
Բաժին 2.3-ում նկարագրվել են R միջավայրում առկա «infotheo» փաթեթի գործառույթները, այդ փաթեթի ուսումնասիրության արդյունքները, ըստ որոնց փաթեթում առկա բոլոր ֆունկցիաները միահոսք են, որը մեծ քանակի տվյալների դեպքում կարող է հանգեցնել արտադրողականության և արդյունավետության նվազեցման:

<u>Երրորդ</u> գլուխը նվիրված է կենսաչափական անվտանգության համակարգերի նկարագրությանը և գաղտնիք գեներացնող համակարգի մոդելի ինֆորմացիոն-տեսական հետազոտությանը:

Բաժին 3.1-ում նկարագրվել է կենսաչափական վավերացման հաղորդակարգը և կատարվել են համապատասխան նշանակումներ: Հայտնի կենսաչափական գաղտնի

[9] E. Haroutunian, "On bounds for E-capacity of DMC," IEEE Transactions on Information Theory, vol. 53, no. 11, pp. 4210-4220, 2007.

վավերացման հաղորդակարգերից մեկը պատկերված է նկար 1-ում: Գրանցման ժամանակ, անհատի կենսաչափական տվյալները վերցվում են, վերլուծվում են և արտահանվում է X^N հաջորդականությունը:



Նկար 1 - Կենսաչափական ապահով վավերացման հաղորդակարգ.

S գաղտնի բանալին կամ ընտրվում է կամ ստեղծվում է X^N հաջորդականությունից: X^N հաջորդականությունից ստեղծվում է նաև M օգնական տվյալը: Գաղտնի բանալին միակողմանի ֆունկցիաների միջոցով կոդավորվում է և պահպանվում է տվյալների բազայում, որպես $f(S)$, ID օգտագործողի համարի, M օգնական տվյալի հետ միասին:

Վավերացման ժամանակ, օգտագործողը ներկայացնում է իր համարը: Նրա կենսաչափական տվյալները կրկին վերցվում են և վերամշակվում, որպես արդյունք ստացվում է Y^N հաջորդականությունը: $\hat{S}$ բանալին մոտարկվում է Y^N-ի և M օգնող տվյալի միջոցով: Մոտարկված $\hat{S}$ բանալին կոդավորվում է նույն միակողմանի ֆունկցիաների միջոցով և համեմատվում է $f(S)$-ի հետ: Օգտագործողին մուտք տրամադրվում է միայն այն դեպքում, երբ $f(S)$-ը հավասար է լինում $f(\hat{S})$-ին:

Բաժին 3.2-ում դիտարկվել է գաղտնի բանալիների գեներացման համար օգտագործվող Ջուելս-Վատենբերգի սխեման[10]: Այն կարելի է նկարագրել հետևյալ կերպ՝

- Գրանցման փուլ:

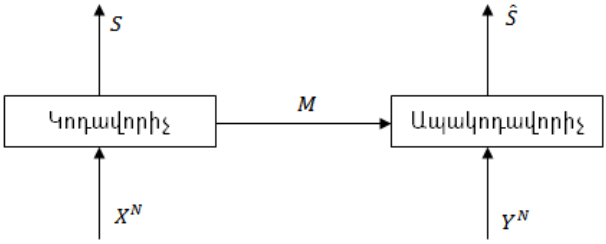
[10] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in Proc. 6th ACM Conf. Computer and Comm. Security, pp. 28-36, 1999.

1. Ընտրել պատահական S բանալի և համապատասխանորեն կառուցել C կոդային բառը, օգտվելով արդեն իսկ ընտրված սխալ ուղղող կոդից:
2. Հաշվարկել և պահպանել $f(S)$-ը և $M=C + X^N$:

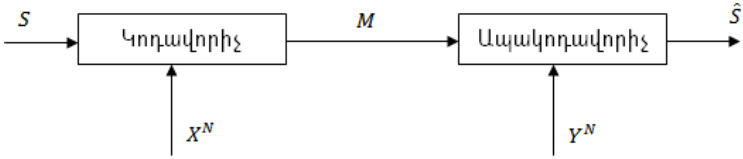
- Նույնականացման փուլ:
  1. Կարդալ Y^N-ը և հաշվարկել $C' = Y^N + M = Y^N + X^N + C = C + E$, որտեղ $v = Y^N + X^N$;
  2. Վերծանել C'-ը և վերականգնել $\hat{S}$-ը: Ճիշտ ապակոդավորման արդյունքում պետք է ստացվի $\hat{S} = S$:

Հաշվարկել $f(\hat{S})$ և համեմատել $f(S)$-ի հետ: Եթե նույնն են ընդունել, հակառակ դեպքում մերժել:

Բաժին 3.3-ում նկարագրվել են գաղտնի բանալի բաշխող կենսաչափական մոդելները և ձևակերպվել են այդ մոդելների բնութագրիչները, կատարվել են համապատասխան նշանակումներ: Ընդհանուր առմամբ, գաղտնի բանալի բաշխող մոդելները կարելի է խմբավորել 2 դասի մեջ՝ մոդելներ գեներացվող գաղտնի բանալիով (տես նկար 2) և մոդելներ ընտրվող գաղտնիքով (տես նկար 3):



Նկար 2 - Գաղտնի բանալի գեներացնող կենսաչափական մոդել.



Նկար 3 - Ընտրվող գաղտնի բանալիով կենսաչափական մոդել.

Երկու մոդելներում էլ S-ը կամայական գեներացված կամ ընտրված գաղտնի բանալի է, X^N-ը գրանցման ժամանակ անհատի կենսաչափական հաջորդականությունն է, իսկ Y^N-ը վավերացման հաջորդականությունն է, երկու հաջորդականություններն էլ ունեն N երկարություն: M-ը օգնական տվյալն է, իսկ $\hat{S}$-ը մոտարկված գաղտնի բանալին: Կոդավորիչը և ապակոդավորիչը միացնող կապուղին ենթադրվում է, որ հասանելի է բոլորին: Ենթադրվում է, որ պասսիվ հարձակումներ հնարավոր են, օրինակ հարձակվողը կարող է տեսնել կապուղում բոլորին հասանելի ինֆորմացիան, բայց այն փոփոխել չի կարող: Ինֆորմացիայի արտահոսքը բնութագրվում է փոխադարձ

ինֆորմացիայի և գաղտնիքի երկարության էնտրոպիայի միջոցով: Նշված երկու դասի մոդելներն էլ պետք է բավարարեն հետևյալ պայմաններին՝

$$\Pr\{S \neq \hat{S}\} \approx 0 \qquad \text{(հուսալիություն)},$$

$$\frac{1}{N}H(S) \approx \frac{1}{N}\log_2|S| \quad \text{(գաղտնիքի հավասարաչափություն)},$$

$\frac{1}{N}H(S)$ որքան հնարավոր է մեծ (գաղտնիքի արագություն),

$$\frac{1}{N}I(S \wedge M) \approx 0 \quad \text{(գաղտնիության արտահոսք)},$$

$\frac{1}{N}I(X^N \wedge M)$ որքան հնարավոր է փոքր (ինքնության արտահոսք):

M-երը անվանում են օգնական տվյալներ, դրանք պետք է աննշան ինֆորմացիա պարունակեն գաղտնի բանալու մասին: Ցույց է տրվել, որ նշված մոդելում, անկախ և հավասարաչափ բաշխման դեպքում ամենամեծ գաղտնիքի արագությունը հավասար է X^N և Y^N ուսումնասիրվող հաջորդականությունների միջև եղած $I(X^N \wedge Y^N)$ փոխադարձ ինֆորմացիային:

Դիտարկենք գաղտնիք գեներացնող կենսաչափական մոդելը (տես. նկար 2), որը հիմնված է $\{Q(x,y), x \in \mathcal{X}, y \in \mathcal{Y}\}$ բաշխումով կենսաչափական աղբյուրի վրա: Այս աղբյուրը $\mathcal{X}$ վերջավոր այբուբենից արտադրում է $\mathbf{x} \equiv x^N = (x_1, x_2, \dots, x_N)$ առաջին հաջորդականությունը, որը բաղկացած է N քանակի ազդանշաններից, և $\mathcal{Y}$ վերջավոր այբուբենից արտադրում է $\mathbf{y} \equiv y^N = (y_1, y_2, \dots, y_N)$ երկրորդ հաջորդականությունը, բաղկացած ևս N քանակի ազդանշաններից: Առաջին հաջորդականությունը հայտնի է գրանցման հաջորդականություն անունով, իսկ երկրորդը՝ վավերականացման հաջորդականություն անունով: Ընդհանուր առմամբ X^N և Y^N կենսաչափական հաջորդականությունները միմյանցից անկախ չեն: Ավելին, Y^N երկրորդ հաջորդականությունը X^N-ի աղմուկով աղավաղված տարբերակն է: Կատարենք հետևյալ նշանակումը՝

$$Q(x,y) = Q_1(x)Q_2(y|x),\ x \in \mathcal{X},\ y \in \mathcal{Y}:$$

Մենք ենթադրում ենք, որ

$$Q^N(\mathbf{x},\mathbf{y}) = \prod_{n=1}^{N} Q(x_n, y_n):$$

Դիտարկենք կոդավորիչ, որը հետազոտում է X^N գրանցման հաջորդականությունը: Այս հաջորդականությունից գաղտնիք գեներացնող և բաշխող կենսաչափական մոդելում կոդավորիչը գեներացնում է $S \in \{1,2,\dots,|\mathcal{S}|\}$ գաղտնիքը և բոլորին հասանելի $M \in \{1,2,\dots,|\mathcal{M}|\}$ օգնող տվյալը: Սա նշանակում է, որ

$$f(X^N) = (S,M),$$

որտեղ $f(\cdot)$-ով նշանակում ենք կոդավորիչի ֆունկցիան: Օգնող տվյալը ուղարկվում է ապակոդավորիչին: Ապակոդավորիչը իր հերթին հետազոտում է վավերականացման Y^N հաջորդականությունը և ստեղծում է S գաղտնիքի $\hat{S}$ մոտարկումը, օգտագործելով ստացած M օգնող տվյալը, հետևաբար

$$g(Y^N, M) = \hat{S},$$

որտեղ $g(\cdot,\cdot)$-ով նշանակում են ապակոդավորիչի ֆունկցիան: Մենք ենթադրում ենք, որ հարձակվողն ունի մուտք դեպի կապուղի, այսինքն նա կարող է տեսնել ամբողջ ինֆորմացիան, բայց չի կարող կատարել փոփոխություններ:

Կենսաչափական գաղտնիություն ապահովող համակարգի կարևորագույն բնութագրիչներն են գաղտնիքի ծավալը և ինքնության արտահոսքը, այսինքն կենսաչափական ինֆորմացիայի քանակը, որն օգնող տվյալների միջոցով կարող է արտահոսվել դուրս: Բնական է պահանջել, որպեսզի ինքնության արտահոսքը լինի փոքր: Ավելին, գաղտնի բանալու ծավալը պետք է լինի մեծ, որպեսզի փոքրացնի գաղտնիքը գուշակելու հավանականությունը: Երկու կողմերի (կոդավորիչի և ապակոդավորիչի) նպատակն է փոքրացնել սխալի հավանականությունը, այսինքն մոտարկված $\hat{S}$ գաղտնիքը գեներացված S գաղտնիքին հավասար չլինելու հավանականությունը մոտեցնել զերոյի:

Սահմանում. Գաղտնի բանալու R արագությունը, երբ $R \geq 0$, կոչվում է <u>*հասանելի*</u>, եթե բոլոր $\delta > 0$ համար և բավականաչափ մեծ N-ի համար, գոյություն ունի կոդ, որը բավարարում է հետևյալ պայմաններին՝

$$\Pr\{S \neq \hat{S}\} \leq \delta,$$
$$\frac{1}{N}H(S) + \delta \geq \frac{1}{N}\log_2|S| \geq R - \delta,$$
$$\frac{1}{N}I(S \wedge M) \leq \delta:$$

Հայտնի էր, որ <u>*գաղտնիք գեներացնող կենսաչափական մոդելի ունակությունը*</u> հավասար է $I(X \wedge Y)$ փոխադարձ ինֆորմացիայի մաքսիմալ արժեքին:

Բաժին 3.4-ում ներմուծվել է նոր E –հասանելի գաղտնիքի արագության գաղափարը գաղտնիք գեներացնող կենսաչափական մոդելի համար: Այդ նոր բնութագրիչի ամենամեծ արժեքի համար դուրս են բերվել ստորին և վերին գնահատականները: Երբ $E \to 0$, ստացված ստորին և վերին գնահատականները համընկնում են և հավասարվում ամենամեծ հասանելի գաղտնիքի արագության հետ:

Սահմանում. Գաղտնի բանալու <u>$R(E)$ *արագությունը*</u>, երբ $R(E) \geq 0$, կոչվում է <u>*E –հասանելի*</u>, եթե բոլոր $\delta > 0$, $E > 0$ համար և բավականաչափ մեծ N-ի համար, գոյություն ունի կոդ, որը բավարարում է հետևյալ պայմաններին՝

$$\Pr\{S \neq \hat{S}\} \leq 2^{-N(E-\delta)},$$
$$\frac{1}{N}H(S) + \delta \geq \frac{1}{N}log_2|S| \geq R(E) - \delta,$$
$$\frac{1}{N}I(S \wedge M) \leq \delta.$$

Արդյունքների ձևակերպման համար մենք օգտագործում ենք հետևյալ բաշխումները՝

$Q_1 = \{Q_1(x), x \in \mathcal{X}\}$, $Q_2 = \{Q_2(y|x), y \in \mathcal{Y}, x \in \mathcal{X}\}$,
$P_1 = \{P_1(x), x \in \mathcal{X}\}$, $P_2 = \{P_2(y|x), y \in \mathcal{Y}, x \in \mathcal{X}\}$,
$Q = \{Q(x,y), x \in \mathcal{X}, y \in \mathcal{Y}\}$,
$P = \{P(x,y), x \in \mathcal{X}, y \in \mathcal{Y}\}$.

Ատենախոսության գլխավոր արդյունքներից է հետևյալ թեորեմը:

Թեորեմ. Գաղտնիք գեներացնող կենսաչափական մոդելում ամենամեծ E –հասանելի գաղտնի բանալու $R(E)$ արագության ստորին գնահատականն է՝

$$R_r(E) = \min_{P:D(P||Q)\le E} |I_P(X \wedge Y) + D(P||Q) - E|^+$$

և վերին գնահատականն է՝

$$R_{sp}(E) = \min_{P:D(P||Q)\le E} I_P(X \wedge Y).$$

Հետևանք. Երբ $E \to 0$, մեր ստացած ստորին և վերին գնահատականները համընկնում են և հավասար են ամենամեծ հասանելի գաղտնիքի արագության հետ՝

$$\lim_{E\to 0} R_r(E) = \lim_{E\to 0} R_{sp}(E) = I_Q(X \wedge Y).$$

Բաժին 3.5-ում տրվել է վերը բերված թեորեմի ապացուցումը: Ապացուցումը հիմնված է կազմերի մեթոդի վրա: Ապացուցումը բացկացած է երկու մասից: Առաջին մասում տրվում է վերին գնահատականի ապացուցումը, իսկ երկրորդ մասը նվիրված է թեորեմի հասանելիությանը:

Բաժին 3.6-ում ձևակերպվել է գաղտնիք գեներացնող կենսաչափական մոդելի ինքնության արտահոսքի վերաբերյալ պնդում և տրվել է այդ պնդման ապացույցը:

Պնդում. Գաղտնիք գներացնող և բաշխող կենսաչափական մոդելում E –հասանելի գաղտնիքի ամենամեծ արագության դեպքում ինքնության արտահոսքը՝

$$\frac{1}{N} I_{Q_1}(M \wedge X^N) = \max_{P:D(P||Q)\le E} (H_P(X|Y) + D(P||Q) - E).$$

Ատենախոսության ***չորրորդ*** գլուխը նվիրված է R ծրագրավորման միջավայրի համար մշակված նոր «Advanced Inftheo» մոդուլի նկարագրությանը, առկա գործառույթներին և նրանում առկա տեխնիկական լուծումներին:

Բաժին 4.1-ում նկարագրվել են «Advanced Inftheo» մոդուլը և նրանում առկա գործառույթները, զուգահեռացման հնարավորությունները և առանձնահատկությունները: Մոդուլը ստեղծվել է C++ լեզվով, այն ընդգրկում է ֆունկցիաներ, որոնցով կարելի է հաշվել ինֆորմացիայի տեսության տարբեր տիպի մեծություններ, ինչպիսիք են՝ արագություն-հուսալիություն ֆունկցիայի ստորին և վերին գնահատականները, փոխադարձ ինֆորմացիան, պայմանական փոխադարձ ինֆորմացիան, Կուլբակի-Լեյբլերի տարամիտությունը և այլն: Մոդուլը օգտագործողին տալիս է երեք տիպի զուգահեռացման հնարավորութուններ, դրանք են՝

- զուգահեռացում տեղային պրոցեսորի միջոցով (CPU տիպ),
- զուգահեռացում գրաֆիկական քարտի միջոցով (GPU տիպ),
- զուգահեռացում կլաստերի միջոցով (MPI տիպ):

Բաժին 4.2-ում նկարագրվել են R միջավայրում առկա սահմանափակումները, որոնք հաղթահարվել են «Advanced Inftheo» փաթեթում: Այդ սահմանափակումներից են

հաշվարկի կատարման ցածր արագությունը և ոչ արդյունավետ հիշողության բաշխումը:

Բաժին 4.3-ում նկարագրվել են «Advanced Inftheo» փաթեթում առկա տեխնիկական լուծումները: «Advanced Inftheo»-ն օգտագործում է բազմահոսքայնությունը ավելի բարձր արտադրողականության հասնելու համար [11]: Բազմահոսքայնության օգտագործումը իր հերթին ենթադրում է, որ նոր սահմանափակումներ են ի հայտ գալիս զուգահեռ աշխատող հոսքերի գործարկման ժամանակ: Փականքները գլխավոր տեխնիկական միջոցներից են, որոնք կիրառվում են «Advanced Inftheo»-ում: Փականքներին հաճախ անվանում են նաև մոնիտորներ, ճգնաժամային հատվածներ, մուտեքսներ կամ սեմաֆորներ, բայց անկախ անվանումից դրանք կատարում են միևնույն գործառույթը: Փակուղիներից խուսափելու համար «Advanced Inftheo»-ում ամեն մի փականքի վերագրվել է աստիճան և ֆունկցիաները նախագծվել են այնպես, որ հոսքերը ազատում են փականքները միայն նվազման կարգով ըստ աստիճանի: Այս եղանակը ստեղծում է փականքներ ներառող ցիկլեր և հետևաբար փակուղին դարձնում անհնարին: Հաշվողական պրոցեսը արագացնելու համար մոդուլում առկա է կլաստերի վրա ֆունկցիաների իրագործման հնարավորություն: Այս հնարավորության իրականացման գլխավոր հատվածը OpenMPI գրադարանի կցորդումն էր մոդուլին: OpenMPI գրադարանի օգտագործումը պայմանավորված էր այն հանգամանքով, որ նրա ճշգրտորեն նախագծված մոդուլային կոմպոնենտի ճարտարապետությունը[12] բաժանում է «Advanced Inftheo»-ի ֆունկցիաները նեղ խմբավորված փոքր մոդուլների, որոնք կարող են անկախորեն իրականացվել և փոփոխվել: OpenMPI-ի մոդուլային կոմպոնենտի ճարտարապետությունը կառավարում է բաղադրիչների հարթակները և տալիս է նրան որոշակի ծառայություններ, ինչպիսիք են օրինակ, գործարկվող բարձր աստիճանի ծրագրից պարամետրեր ընդունելու և ցածր մակարդակի առանձին բաղադրիչների փոխանցելու հնարավորությունը: Ամեն մի կոմպոնենտի հարթակ նախատեսված է կոնկրետ առաջադրանքի համար:

Բաժին 4.4-ում նկարագրվել են կոնֆիգուրացվող պարամետրերով գաղտնիք գեներացնող և բաշխող կենսաչափական համակարգի կառուցման տեխնիկական մանրամասները, որի հիմքում ընկած է անձի մատնահետքը որպես կենսաչափական եզակի բնութագրիչ: Առաջարկվող գաղտնիք գեներացնող և բաշխող կենսաչափական համակարգի տեխնիկական հիմքը կառուցվելու է «Raspberry Pi» միահարթակ, փոքր չափի համակարգչի և TCS4K կենսաչափական սենսորի օգնությամբ:

Ատենախոսության ***<u>հինգերորդ</u>*** գլուխը նվիրված է «Advanced Inftheo» մոդուլի արագագործության վերլուծություններին:

---

[11] C. Hughes and T.Hughes, Professional Multicore Programming: Design and Implementation for C++ Developers. Birmingham, UK: UK: Wrox Press Ltd., 2008.

[12] G. Fagg, G. Bosilca and E. Gabriel, "Open MPI: Goals, concept, and design of a next generation MPI implementation," in *Proceedings of 11th European PVM/MPI Users' Group Meeting*, Budapest, Hungary, pp. 97-104, 2004.

Բաժին 5.1-ում նկարագրվել են «Advanced Inftheo» մոդուլի փորձարկման արդյունքները գաղտնիք գեներացնող կենսաչափական համակարգի գաղտնիքի E-հասանելի արագության գնահատականների որոշման օրինակի վրա:

Բաժին 5.2-ում ներկայացվում են «Advanced Inftheo» մոդուլի արտադրողականության վերլուծությունների արդյունքները և ձևակերպվում են ընդհանուր խորհուրդներ հոսքերի օպտիմալ քանակի ընտրման համար, ինչպես նաև համեմատվում են երկու մոդուլների՝ «Infotheo»-ի և «Advanced Inftheo»-ի միանման ֆունկցիաների արտադրողականությունը:

Ատենախոսությունը ամփոփված է եզրակացությունով:

Աշխատանքի հիմնական արդյունքները

Աշխատանքում ստացվել են հետևյալ հիմնական արդյունքները՝

1. Գաղտնիք գեներացնող կենսաչափական համակարգի համար ներմուծվել է գաղտնիքի E –հասանելի արագություն նոր հասկացությունը, որն ընդհանրացնում է գաղտնիքի հասանելի արագության գաղափարը [1]:
2. Կառուցվել են գաղտնիք գեներացնող կենսաչափական համակարգի ամենամեծ գաղտնիքի E –հասանելի արագության վերին և ստորին գնահատականները: Երբ $E \to 0$, դրանք համընկնում են և հավասար են գաղտնի բանալու հասանելի արագության մեծագույն արժեքին [4]:
3. Ներկայացվել է փոփոխելի հուսալիության աստիճան ունեցող գաղտնիք գեներացնող կենսաչափական համակարգի մոդել [2]:
4. Մշակվել է նոր փաթեթ R միջավայրում ինֆորմացիայի տեսության բարդ ֆունկցիաների հաշվարկի համար [3, 5]:
5. Մշակված մոդուլը առաջարկում է հաշվարկի երեք տեսակի զուգահեռացման հնարավորություն և թույլ է տալիս փոփոխել զուգահեռացման տեսակը հաշվարկի ընթացքում [6]:

Հրապարակված աշխատությունների ցանկը

1. M. Haroutunian, N. Pahlevanyan, “Information theoretical analysis of biometric secret key sharing model,” Transactions of IIAP of NAS of RA, Mathematical Problems of Computer Science, vol. 42, pp. 17-27, 2014.
2. M. Haroutunian, N. Pahlevanyan, “Technical aspects of biometric secret sharing system construction,” Scientific Proceedings of GSPI, Vol. A, Number 1, pp. 116-121, 2014.
3. Н. Пайлеванян “Технологические решения вычислений теоретико-информационных результатов в коммуникационных системах,” Международная научно-техническая конференция студентов и молодых специалистов из стран-

участниц РСС по направлениям «Информационные технологии и системы связи», Сборник тезисов, стр. 96-99, Москва, 2015.

4. M. Haroutunian, N. Pahlevanyan, “Experimentation of Advanced Inftheo module for R on the example of biometric generated secret key sharing system,” International Journal “Information Content and Processing”, Vol. 2, Number 1, pp. 62-70, 2015.
5. N. Pahlevanyan, M. Haroutunian, “Technical solutions of developing Advanced Inftheo new module for R,” Proceedings of the 10th International Conference on Computer Science and Information Technologies, Yerevan, Armenia, pp. 306-309, 2015.
6. N. Pahlevanyan, M. Haroutunian, “Results of performance analysis of Advanced Inftheo new package for R,” Transactions of IIAP of NAS of RA, Mathematical Problems of Computer Science, vol. 45, pp. 5-13, 2016.

ИНФОРМАЦИОННО ТЕОРЕТИЧЕСКИЙ АНАЛИЗ БИОМЕТРИЧЕСКОЙ СИСТЕМЫ С ГЕНЕРИРУЕМЫМ СЕКРЕТОМ И СОЗДАНИЕ НОВОГО МОДУЛЯ ДЛЯ СРЕДЫ R

РЕЗЮМЕ

Актуальность исследования. В последние годы, важной частью разработки средств безопасности стали исследования различных систем, основанных на применении биометрических данных.

Биометрические системы секретности основаны на физиологических или поведенческих характеристиках человека, таких как отпечатки пальцев, параметры руки, лица, голоса, особенности сетчатки глаза и т.д. Преимущество биометрических методов секретности заключается в том, что они являются уникальными идентификаторами для человека, которые не могут быть потеряны или фальсифицированы. Но уникальность биометрических данных несет с собой несколько проблем, например, в случае необходимости биометрические данные не могут быть изменены. Эти проблемы предполагают создание и анализ более совершенных систем.

Любая биометрическая система с информационно-теоретической точки зрения может обеспечить секретность до определенного уровня. В зависимости от применения рассматриваются различные модели биометрической аутентификации и генерации секретного ключа. Для каждой модели первой важной задачей является определение достижимой скорости секретного ключа. Второй задачей теории информации является исследование функции надежности. Она представляет собой функцию взаимозависимости основных характеристик (скорость секретного ключа, надежность, вероятность ошибки, скорость утечки) и обобщает понятие максимально достижимой скорости. На практике исследование функции надежности является сложной задачей. Об этом свидетельствует тот факт, что даже для простого дискретного канала без памяти, точная аналитическая форма функции надежности неизвестна, известны только верхняя и нижняя границы этой функции. Полученные теоретические результаты трудно вычисляемы из-за большого количества распределений. Следовательно, разработка программного обеспечения, позволяющего выполнять эти вычисления, является актуальной задачей.

Использование среды R в современном обществе растет очень быстро из-за ряда преимуществ, которые она имеет по сравнению с другими статистическими инструментами. Среда R предназначена для сбора данных, статистической обработки, анализа временных рядов, линейного и нелинейного моделирования, классификации и реализации различных математических расчетов. Среда R может быть использована для расчетов приведенных выше формул, но существующий модуль теории информации в среде R имеет очень ограниченные функциональные возможности.

В этой ситуации актуальна задача разработки более усовершенствованного модуля в среде R для оценки и расчета сложных формул теории информации.

Цель исследования. Для вычислений сложных формул теории информации была поставлена задача - разработать новый модуль для среды R. Проведено информационно-теоретическое исследование биометрической модели распределения сгенерированного секретного ключа.

Научная новизна исследования.

1. Введено новое понятие -достижимой скорости секрета в биометрической модели распределения сгенерированного секретного ключа. Получены верхняя и нижняя границы максимальной E-достижимой скорости секрета. Когда $E \to 0$, пределы верхней и нижней границ -достижимой скорости секрета совпадают с наибольшей достижимой скоростью секретного ключа.
2. Разработан новый модуль среды R для оценки и расчета сложных формул теории информации.
3. Обеспечено три типа параллелизации, которые могут быть изменены в процессе вычислений.
4. Разработаны механизмы для решения проблем, связанных с многопоточностью внутри модуля, например, использование блокировок (locks) и избежание ситуаций взаимных блокировок (deadlock).

Практическая значимость исследования. Разработанный модуль позволяет рассчитать не только энтропию, расстояние Кульбака-Лейблера, среднюю взаимную информацию, но и наименьшее значение этих функций при определенных условиях.

Новый модуль предлагает три типа параллелизации - через локальный процессор (тип CPU), видеокарту (тип GPU) и кластер (тип MPI).

Модуль может быть использован специалистами в области теории информации для расчета ряда сложных функций.

Результаты. В работе получены следующие основные результаты:

- Введено новое понятие -достижимой скорости секрета для биометрической модели с генерируемым секретом, которая является обобщением достижимой скорости секрета.
- Получены верхняя и нижняя границы максимальной -достижимой скорости секрета для биометрической модели с генерируемым секретом. Когда $E \to 0$, пределы верхней и нижней границ -достижимой скорости секрета совпадают с наибольшей достижимой скоростью секретного ключа [4].
- Представлена биометрическая система распределения с генерируемым секретом с настраиваемым уровнем надежности [2].
- Разработан новый модуль в среде R для оценки и расчета сложных формул теории информации [3, 5].
- Разработана опция параллелизации в вычислениях внутри модуля, которая дает возможность изменений типа параллелизации в процессе вычислений [6].

INFORMATION THEORETICAL ANALYSIS OF BIOMETRIC GENERATED SECRET KEY SHARING SYSTEM AND DEVELOPMENT OF NEW PACKAGE FOR R ENVIRONMENT

SUMMARY

The relevance of the research.

In recent years, the exploration of various systems based on biometric data is becoming an important part of the security sector. Biometric secrecy systems are based on the person's physiological or behavioral characteristics such as fingerprints, hand geometry, facial, vocal, iris, retinal features etc. The advantage of biometric data over traditional secrecy methods is the fact that they are unique identifiers for the individual, which could not be lost or falsified. But the uniqueness of biometric data causes problems, such as in case of necessity biometric data can not be changed. These problems suggest analyses of more complex systems.

Any biometric secrecy system can be information-theoretically secured up to a certain level. Depending on applications different models of biometric authentication and secret generation are considered. The first important task for each model is to determine the achievable secret key rate. The next task of Information Theory is investigation of reliability function. It represents coherent function of main characteristics (secret key rate, reliability, error probability, leakage rate) and summarizes the idea of the largest achievable rate. The evidence of the complexity is the fact that analytic form of the function is unknown even for the simple Discrete Memoryless Channel, only the upper and lower bounds are known. Hence, the development of software that can perform these calculations is a relevant task.

The usage of R environment in modern society is growing very quickly due to a number of advantages it has compared to other statistical tools. R environment is designed for data collection, statistical processing, time series analysis, linear and nonlinear modeling, data classification, implementation of various mathematical calculations and construction of corresponding graphs. R can contribute to calculations of the above formulas, but the existing package of Information Theory has very limited functionality.

In this situation it is important to develop an improved package for R environment for estimation and computation of complex formulas of Information Theory.

The aim of the research.

To develop a new package for R environment for performing estimation and computation of complicated formulas of Information Theory. For implementation of the package it was tasked to carry out the information-theoretical investigation of coherent functions of the main characteristics (secret key rate, reliability, error probability, leakage rate) of the biometric secret key sharing model.

Scientific novelty of the research.

1. A new concept of E-achievable secret key rate for biometric generated secret key sharing model is introduced and the expressions for the lower and upper

bounds of the largest rate are obtained. When $E \to 0$, the limits of bounds coincide and are equal the largest achievable secret key rate.

2. The new software package for R environment for estimation and computation of complex formulas of Information Theory is created.
3. Three types of parallelization inside the package, that can be changed during computation process.
4. Mechanisms are developed for solving issues related to multithreading inside package, such as usage of locks and avoiding deadlock situations.

Practical significance of the research:

The developed package allows calculation of not only entropy, Kullback-Leibler (divergence) distance, the average mutual information, but also the smallest values of these functions under certain conditions.

The new package offers three types of parallelization - through the local processor (CPU type), graphic card (GPU type) and cluster (MPI type).

The package can be used by specialists in Information Theory for calculation of a number of sophisticated functions.

Results of the research.

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

- A new concept of E-achievable secret key rate for biometric generated secret key sharing model has been introduced, which is the generalization of the secret key rate studied by Ignatenko and Willems [1].
- For biometric generated secret key sharing model the lower and upper bounds of the largest E-achievable secret key rate have been obtained. When $E \to 0$, the limits of bounds coincide and are equal to the largest achievable secret key rate [4].
- The biometric secret sharing system with configurable reliability level has been presented [2].
- The new package for R environment for estimation and computation of complex formulas of Information Theory has been developed [3, 5].
- An option for using of the three types of parallelization inside package has been developed, which allows change of parallelization type during computation process [6].

Ծավալը՝ 20 էջ: Տպաքանակը՝ 100:
ՀՀ ԳԱԱ ԻԱՊԻ կոմպյուտերային պոլիգրաֆիայի լաբորատորիա:
Երևան, Պ. Սևակի 1