НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
Навчально-науковий фізико-технічний інститут
Кафедра математичних методів захисту інформації

«На правах рукопису»
УДК 519.72

«До захисту допущено»
В.о. завідувача кафедри
_____ Сергій ЯКОВЛЄВ
«__» _____ 2023 р.

Магістерська дисертація

на здобуття ступеня магістра
за освітньо-професійною програмою
«Математичні методи криптографічного захисту інформації»

зі спеціальності: 113 Прикладна математика
на тему: **«Дослідження застосування моделей цінності інформації в криптографії»**

Виконав:
студент II курсу, групи ФІ-12мн
Гетьман Дмитро Олекійович _____

Керівник:
професор, доктор фіз.-мат. наук, доцент
Савчук Михайло Миколайович _____

Рецензент:
посада, степінь, звання
Прізвище Ім'я По-батькові _____

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.

Студент _____

Київ — 2023

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
Навчально-науковий фізико-технічний інститут
Кафедра математичних методів захисту інформації

Рівень вищої освіти — другий (магістерський)
Спеціальність — 113 Прикладна математика,
ОПП «Математичні методи криптографічного захисту інформації»

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
_____ Сергій ЯКОВЛЄВ
«__» _____ 2023 р.

ЗАВДАННЯ
на магістерську дисертацію

Студент: Гетьман Дмитро Олекійович

1. Тема роботи: «*Дослідження застосування моделей цінності інформації в криптографії*», науковий керівник дисертації: професор, доктор фіз.-мат. наук, доцент Савчук Михайло Миколайович,

затверджені наказом по університету №__ від «__» _____ 2023 р.

2. Термін подання студентом роботи: «__» _____ 2023 р.

3. Об'єкт дослідження: інформаційні процеси в системах обробки, застосуавння та захисту інформації.

4. Предмет дослідження: математичні моделі цінності інформації із використаням в крипторафічному захисті.

5. Перелік завдань: огляд опублікованих джерел за тематикою дослідження; дослідження застосування теорії цінності інформації в криптографії; узагальнення обранної моделі цінності інформації на широких спектр розподілів; реалізація узагальненої моделі; аналіз отриманих результатів.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: презентація доповіді.

7. Орієнтовний перелік публікацій: планується доповідь на всеукраїнській конференції.

8. Дата видачі завдання: 10 вересня 2022 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання	Примітка
1	Узгодження теми роботи із науковим керівником	01-15 вересня 2022 р.	Виконано
2	Огляд опублікованих джерел за тематикою дослідження	Вересень-жовтень 2022 р.	Виконано
3	Дослідження наявних моделей цінності інформації та вибір однієї із них для подальшого узагальнення на широкий спектр рзподілів	Листопад-грудень 2022 р.	Виконано
4	Програмна реалізація узагальненної моделі та створення веб-застосунку	Січень-березень 2023 р.	Виконано
5	Аналіз отриманих результатів та їх оформлення	Квітень-травень 2023 р.	Виконано

Студент _____ Дмитро Гетьман

Керівник _____ Михайло Савчук

РЕФЕРАТ

Кваліфікаційна робота містить: 43 стор., 4 рисунки, 3 таблиць, 10 джерел.

Метою кваліфікаційної роботи є дослідження застосування моделей цінності інформації в криптографії. Завданням кваліфікаційної роботи є узагальнення обраної моделі цінності інформації, її опис, програмна реалізація у вигляді веб-застосунку; аналіз та порівняння отриманих цінностей інформації в різних моделях та з різними вхідними параметрами.

Об'єктом дослідження є інформаційні процеси в системах обробки, застосуавння та захисту інформації.

Предметом дослідження є математичні моделі цінності інформації із використаням в крипторафічному захисті.

В ході дослідження запропоновано та описано узагальнення однієї моделі цінності інформації. Розкрито застосування теорії цінності інформації в криптографії. Зроблено веб-застосунок, в якому реалізовано чотири варіанта узагальненої моделі цінності інформації. Отримані результати роботи різних варіантів моделі говорять про те, що середня цінність нформації завжди вища при викоритсанні наступних вхідних параметрів моделі: розподіл ймовірностей на всіх можливих станах об'єкта спостереження є нерівноймовірним і користувач обирає найбільш ймовірній стан із повідомлення. Узагальнення моделі на широкий спектр розподілів та використання оптимальної стратегії при знаходженні цінності інформації дозволяє застосувавати цю модель на більш широкий перелік практичних задач та отримувати кращі результати.

КЛЮЧОВІ СЛОВА: ІНФОРМАЦІЯ, ЦІННІСТЬ ІНФОРМАЦІЇ, МАТЕМАТИЧНІ МОДЕЛІ ЦІННОСТІ ІНФОРМАЦІЇ, ІНФОРМАЦІЙНІ РИЗИКИ, ДЕЗІНФОРМАЦІЯ, КРИПТОГРАФІЧНИЙ ЗАХИСТ ІНФОРМАЦІЇ, ОЦІНКА СТІЙКОСТІ КРИПТОСИСТЕМ.

ABSTRACT

The qualification work contains: 43 p., 4 drawings, 3 tables, 10 sources.

The purpose of the qualification work is to research the application of information value models in cryptography. The task of the qualification work is the generalization of the chosen model of the value of information, its description, program implementation in the form of a web application; analysis and comparison of the obtained information values in different models and with different parameters.

The object of research is information processes in the systems of information processing, application and protection.

The subject of research is mathematical models of the value of information with use in cryptographic protection.

During the research, a generalization of the specific information value model was proposed and described. The application of the theory of the value of information in cryptography is disclosed. A web application was made, in which four variants of the generalized model of the value of information were implemented. The obtained results of the work of different versions of the model indicate that the average value of information is always higher when using the following input parameters of the model: the distribution of probabilities on all possible states of the object of observation is unequal and the user chooses the most probable state from the message. The generalization of the model to a wide range of distributions and the use of an optimal strategy when finding the value of information allows you to apply this model to a wider list of practical problems and obtain better results.

KEY WORDS: INFORMATION, INFORMATION VALUE, MATHEMATICAL MODELS OF INFORMATION VALUE, INFORMATION RISKS, DISINFORMATION, CRYPTOGRAPHIC PROTECTION OF INFORMATION, ASSESSMENT OF STABILITY OF CRYPTOSYSTEMS.

ЗМІСТ

ВСТУП

Актуальність дослідження. Актуальність дослідження використання теорії цінності в криптографії полягає в наступних аспектах:

1) Забезпечення конфіденційності даних: Криптографія використовується для захисту конфіденційності даних та забезпечення приватності. Використання теорії цінності може допомогти визначити, які саме дані є найціннішими з точки зору потенційного зловмисника. Це дозволить криптографічним системам надати більший рівень захисту для цих даних, забезпечуючи їх цілісність та недоступність для несанкціонованого доступу.

2) Викриття вразливостей: Дослідження використання теорії цінності може допомогти виявити вразливості в криптографічних системах, що базуються на неправильних оцінках цінності даних. Шляхом визначення ключових місць і їх прихованих зв'язків можна ідентифікувати вразливі точки, які можуть бути використані для атак на систему.

3) Оптимізація ресурсів: Використання теорії цінності може допомогти в розподілі обмежених ресурсів в криптографічних системах. Шляхом визначення цінності різних даних та розробки ефективних алгоритмів і стратегій, можна забезпечити оптимальне використання обчислювальних ресурсів, енергії та інших обмежених складових.

4) Розвиток нових методів шифрування: Теорія цінності може служити основою для розробки нових методів шифрування та криптографічних протоколів. Використання цінності даних для визначення ключових параметрів шифрування може призвести до створення більш ефективних та надійних криптографічних методів, які враховують конкретні потреби та цінності користувачів.

Метою дослідження є розкриття застосування теорії цінності

інформації в криптографії. Для досягнення мети необхідної розв'язати **задачу дослідження**, яка полягає в узагальнені однієї моделі цінності інформації на широкий спектр розподілів та подальшого аналізу отриманих результатів її роботи. Для розв'язання задачі необхідно вирішити такі завдання:

- провести огляд опублікованих джерел за тематикою дослідження;

- обрати одну модель цінності інформації для подальшого узагальнення;

- навести формальний опис узагальненої моделі цінності інформації;

- програмно реалізувати цю модель та дослідити результати її роботи.

Об'єктом дослідження є інформаційні процеси в системах обробки, застосуавння та захисту інформації.

Предметом дослідження є математичні моделі цінності інформації із використаням в крипторафічному захисті.

При розв'язанні поставлених завдань використовувались такі *методи дослідження*: методи теорії ймовірності, методи алгебри, комбінаторного аналізу, методи оцінок ризиків та цінності інфромації, статистичне моделювання.

Наукова новизна отриманих результатів полягає у вперше запропонованому узагальненю однієї моделі цінності інформації на широкий спектр розподілів в системі передачі інформації в умовах невизначеності. Вперше експерементально досліджено роботу даної моделі та проаналізовано результати.

Практичне значення результатів полягає в удосконалені обраної моделі цінності інформації шляхом її узагальнення на широкий спектр розподілів. Таке удосканалення дозволяє використовувати дану модель в більш широкому спектрі практичних задач та збільшити одержувані доходи та цінність інформації.

1 НЕОБХІДНІ ТЕОРЕТИЧНІ ВІДОМОСТІ ПРО ТЕОРІЮ ЦІННОСТІ ІНФОРМАЦІЇ, МОДЕЛІ ЦІННОСТІ ІНОРМАЦІЇ, МОДЕЛЬ ЦІННОСТІ ІНФОРМАЦІЇ ЗА Г. П. ШАНКІНОМ, ЗАСТОСУВАННЯ ТЕОРІЇ ЦІННОСТІ ІНФОРМАЦІЇ В КРИПТОГРАФІЇ

В данному розділі наведені необхідні теоретичні відомості щодо теорії цінності інформації та моделей цінності інформації. Для подальшого дослідження та узагальнення обрана одна модель цінності інформації - модель цінності інформації за Г. П. Шанкіном, наведено опис цієї моделі. Розкриті можливі застосування теорії цінності інформації в криптографії.

1.1 Історична довідка

Теорія цінності інформації має багато історичних корінь, які відображають історію розвитку людської цивілізації і комунікаційних технологій в цілому.

Історія теорії цінності інформації починається з розвитку писемності та поширення книгодрукування в середньовічній Європі. У цей період люди почали збирати та зберігати інформацію, що вважалася цінною. З появою промислової революції та зростанням кількості інформації, яка була доступна, виникла необхідність у класифікації та організації інформації.

У середині XX століття вчені почали активно досліджувати інформацію як економічний ресурс. У 1945 році економіст Фрідріх Гайдер вперше ввів поняття «економіки інформації» і запропонував розглядати інформацію як товар [1].

У 1957 році економіст Герберт Саймон ввів поняття «обмеженості розуміння», яке відображало те, що людина не може обробляти нескінченну кількість інформації і обмежена в своїх здібностях зрозуміти й зберегти інформацію [2].

У 1980-х роках теорія цінності інформації стала досить популярною в інформаційних науках і бібліотечній справі. Вчені розглядали цінність інформації як ключовий фактор в процесі прийняття рішень і вважали, що інформація може використовуватися як інструмент для збільшення ефективності діяльності.

У наш час теорія цінності інформації залишається важливою для розуміння того, як люди отримують, обробляють та використовують інформацію. Цінність інформації може визначатися різними чинниками, такими як актуальність, достовірність, доступність, контекст та специфічні потреби користувача.

З розвитком інформаційних технологій теорія цінності інформації також змінювалася. Зокрема, з'явилися нові методи та техніки збору, аналізу та інтерпретації даних, що сприяють підвищенню рівня точності та ефективності прийняття рішень.

Одним із актуальних напрямків в теорії цінності інформації є вивчення її впливу на людське здоров'я та психіку. Сучасні дослідження показують, що перенавантаження інформацією може спричинити стрес, зниження концентрації та інші негативні наслідки. Також важливим напрямком є дослідження впливу інформації на соціальну сферу, зокрема на політику, культуру та етику. Деякі вчені вважають, що інформація може впливати на формування світогляду, цінностей та поведінки людей.

У цілому, теорія цінності інформації є важливою складовою сучасного інформаційного суспільства та допомагає розуміти, як інформація впливає на людей та як її ефективно використовувати.

1.2 Загальні відомості

Теорія цінності інформації - це концептуальна рамка, яка допомагає зрозуміти, як інформація стає цінною для людей та організацій і як вона використовується для досягнення різних цілей.

Згідно з теорією цінності інформації, інформація має цінність для людей в тому випадку, коли вона може бути використана для досягнення певної мети або цілі. Ця цінність може бути різною для різних людей, в залежності від їх потреб та інтересів.

Крім того, теорія цінності інформації включає в себе концепцію «інформаційного шуму» - інформації, яка не має цінності для користувача і може заважати досягненню мети. Це може бути непотрібна, неточна або надлишкова інформація. Таку інформацію використовують під час дезінформації.

Основна ідея теорії цінності інформації полягає в тому, що краще використовувати інформацію як засіб досягнення певної мети, а не просто збирати її без мети. Це допомагає забезпечити ефективне використання інформації та зменшити кількість «інформаційного шуму», який може заважати досягненню цілей.

Теорія цінності інформації має важливе значення для бізнесу та організацій, оскільки допомагає зрозуміти, як інформація може бути використана для досягнення стратегічних цілей та підвищення конкурентоспроможності. Наприклад, бізнес може зосередитися на зборі та аналізі інформації про потреби та інтереси своїх клієнтів, щоб створювати більш ефективні маркетингові кампанії та продукти.

Також теорія цінності інформації відіграє важливу роль в контексті інформаційної безпеки. Згідно з цією теорією, цінність інформації залежить від її конфіденційності, цілісності та доступності. Інформація може бути цінною тільки тоді, коли вона захищена від несанкціонованого доступу, не змінена та доступна за необхідності.

Наведемо кілька загальних визначень цінності інформації:

1) Цінність інформації проявляється у тому випадку, якщо вона сприяє досягненню мети, що стоїть перед споживачем [3].

2) Цінність будь-якого інформаційного джерела визначається як різниця між доходами від двох оптимальних стратегій, одна з яких забезпечує свободу вибору різних дій, пов'язаних з використанням інформації, а друга полягає у відсутності такої свободи [4].

3) Цінність інформації - властивість інформації, що визначається її придатністю до практичного застосування у різних галузях цілеспрямованої людської діяльності задля досягнення певної мети [5].

4) Цінність інформації – властивість інформації, що визначається придатністю цієї інформації до практичного використання у різних галузях діяльності людини. Розповсюдження та використання інформації використання призводить до зміни її цінності. З плином часу цінність більшості видів інформації зменшується [6].

Головним питанням практичного застосування поняття цінність інформації є знаходження кількісної оцінки цінності інформації. Ця проблема має давню історію. Її дослідженню присвячено значну кількість публікацій, зокрема роботи К. Шеннона, А. А. Харкевича, Р. Л. Стратанович, М. М. Бонгарда та ін., які стали класичними в цій галузі. Існуюче різноманіття підходів і методів визначення цінності інформації об'єктивно обумовлено існуванням різних видів інформаційних систем, в яких обробляється або циркулює інформація, що оцінюється, безліччю несхожих цілей, для реалізації яких використовується ця інформація, особливостями прикладних завдань, до вирішення яких вона застосовується.

Формально цінність інформації можна визначити наступним чином (роботи Архіпова А.Є. [7-9]):

$$V(I) = \Delta A_{extr}(I) - d(I),$$

де A - показник, що характеризує ступінь успішності виконання

певного завдання, роботи, іншого виду діяльності (цим показником може бути вартість продукції, виготовленої за певний час або з фіксованого обсягу вихідної сировини, виграш, обумовлений вибором вдалого рішення, загальна вартість послуг, що надаються споживачам у певній сфері діяльності, тощо); $d(I)$ - витрати на отримання, обробку та використання інформації I у певному виді діяльності; ΔA - покращення (зростання) показника A за рахунок отриманої інформації I:

$$\Delta A(I) = A(I) - A_0,$$

де A_0 – вихідне значення показника (за відсутності інформації I), $A(I)$ – збільшене (завдяки використанню інформації I) значення показника A.

Виконання умов екстремальності обумовлює зростання показника A до максимально можливого значення A_{extr}:

$$\Delta A_{extr}(I) = A_{extr}(I) - A_0.$$

1.3 Моделі цінності інформації

Модель цінності інформації - це концептуальна модель, яка допомагає оцінити цінність інформації з різних точок зору. Ця модель враховує різні фактори, що визначають, наскільки корисна інформація для конкретної особи або організації.

Модель цінності інформації потрібна для того, щоб допомогти людям та організаціям приймати рішення про те, яку інформацію потрібно збирати, обробляти та зберігати, а яку не потрібно. Вона дозволяє визначити, наскільки важливо зберігати інформацію, яку можна використовувати для підвищення ефективності та збільшення доходів.

Модель цінності інформації також допомагає визначити, які джерела інформації є найбільш цінними для організації. Це дозволяє зосередитися на зборі інформації з тих джерел, які найбільш сприятливо

впливають на бізнес-процеси та допомагають знижувати ризики та збільшувати прибуток.

Крім того, модель цінності інформації допомагає визначити, яка інформація може бути найбільш корисною для різних груп користувачів, таких як менеджери, співробітники, клієнти та інші. Це дозволяє забезпечити більш ефективне використання інформації та збільшити загальну продуктивність та ефективність організації.

Існує кілька моделей цінності інформації, що визначають, яку інформацію вважати цінною та яку - ні. Ось кілька з них:

1) Модель цінності інформації за її практичним застосуванням. Згідно з цією моделлю, інформація вважається цінною, якщо вона може бути використана для розв'язання практичних проблем або досягнення конкретних цілей. Наприклад, інформація про нові методики лікування є цінною для лікарів, які шукають ефективні способи лікування своїх пацієнтів.

2) Модель цінності інформації за її новизною. Інформація вважається цінною, якщо вона нова і унікальна. Це може бути інформація про нові відкриття в науці, нові ринки або нові технології. Така інформація може допомогти людям отримати конкурентну перевагу.

3) Модель цінності інформації за її достовірністю. Інформація вважається цінною, якщо вона є правдивою і може бути підтверджена. Наприклад, інформація про результати наукового дослідження, які були підтверджені декількома незалежними дослідниками, вважається цінною.

4) Модель цінності інформації за її важливістю. Інформація вважається цінною, якщо вона є важливою для особистих, професійних або громадських справ. Наприклад, інформація про загрозу безпеці нації є дуже важливою для правоохоронних органів.

5) Модель цінності інформації за її корисністю. Інформація вважається цінною, якщо вона може бути використана для розвитку знань та вмінь людини або для здійснення певних дій. Наприклад, інформація про те, як виробити нову корисну річ своїми руками, може

бути корисною для людей, які хочуть вдосконалити свої навички виготовлення речей.

6) Модель цінності інформації за її доступністю. Інформація вважається цінною, якщо вона легко доступна та зрозуміла для людей. Наприклад, інформація про те, як покращити здоров'я, є цінною, якщо вона доступна для всіх, хто цікавиться цією темою, і представлена зрозуміло.

7) Модель цінності інформації за її вартістю. Інформація вважається цінною, якщо вона має певну економічну вартість. Це може бути інформація про фінансові ринки або інші види інвестицій, які можуть принести дохід.

Кожна з цих моделей має свої переваги та недоліки, і вибір тієї, яка найбільше підходить для конкретної ситуації, залежить від конкретних потреб та обставин.

1.4 Модель цінності інформації за Г. П. Шанкіном

Далі наведені загальні відомості про модель цінності інформації за Г. П. Шанкіном, ця інформації здебільшого взята із першоджерела [10].

В загальном випадку розглядається наступна інформаційно-аналітична система. Маємо об'єкт спостереження, котрий може знаходитись в одному зі станів кінцевої множини Ω. В фіксований момент часу об'єкт спостереження знаходиться в стані $\omega^* \in \Omega$. Стан об'єкта спостереження вивчає деякий спостерігач. Ці дослідження спостерігача в загальному випадку можуть бути неточними в наступному сенсі. Спостерігач не визначає «істинний» стан, а лише виділяє найбільш ймовірну підмножину $\Omega' \subseteq \Omega$, в якій знаходиться ω^*. Результат своїх досліджень спостерігач формує в деякому повідомленні $x \in X$, де X - це скінченна множина можливих повідомлень. Змістом повідомлення $x \in X$ є множина $\Omega(x) \subseteq \Omega$, в якій на думку спостерігача, знаходиться «істинне» повідомлення ω^*. Результат своїх досліджень $x \in X$ спостерігач

направляє по деякому каналу зв'язку користувачеві повідомлення інформації. Користувач використовує цю інформацію для досягнення деякої цілі, пов'язаної з його апріорним знаннями про стан $\omega^* \in \Omega$.

Слід зазначити, що така структура інформаційно-аналітичної системи використовується досить часто при описі теорії цінності інформації.

Дії користувача описуються наступним чином. Виходячи зі своїх апріорних знань і отриманого повідомлення, користувач прагне виділити найбільш ймовірний стан ($\omega^{'} \in \Omega$), в якому знаходиться об'єкт спостереження. Виділивши цей стан, користувач робить певні дії, ефективність яких може бути виміряна кількісно.

Далі будуть наведені деякі твердження, на основі яких формується вихідна, найбільш проста математична модель цінності інформації.

Твердження 1.

Якщо користувач правильно визначив стан об'єкта спостереження, тобто $\omega^{'} = \omega^*$, то ефективність його дій оцінюється величиною $\alpha > 0$. Якщо ж користувач визначив стан об'єкта спостереження неправильно, тобто $\omega^{'} \neq \omega^*$, то ця ефективність дорівнює $0 \leq \beta < \alpha$. Це припущення справедливо для всіх ω^*, $\omega^{'} \in \Omega$; величини α, β не залежать від конкретних ω^*, $\omega^{'} \in \Omega$, і є константами. Модель цінності інформації, яка використовує *Твердження 1.* будемо називати (α, β)-моделлю. При цьому не обговорюється питання про те, наскільки ефективно сам користувач використовує інформацію, яка до нього надійшла. Припускається, що користувач використовує цю інформацію найбільш ефективним (з його точки зору) чином. Відповідно він і не оцінює цінність отриманої інформації.

Твердження 2.

Спотворення в каналі зв'язку відсутні. Користувач отримує те саме повідомлення $x \in X$, яке відправив спостерігач. Затримки в передачі повідомлення від спостерігача до користувача відсутні, тобто повідомлення передається «миттєво».

Твердження 3.

Користувач і спостерігач однаково розуміють зміст повідомлення $x \in X$. Таким чином, проблема «непорозуміння» (або «неповного розуміння») користувачем сенсу переданого йому повідомлення спостерігачем повідомлення знімається. Має місце повне взаєморозуміння спостерігача і користувача.

Твердження 4.

Введемо позначення:

$$X^+ = \{x \in X \mid \omega^* \in \Omega(x)\},$$

$$X^- = X \setminus X^+.$$

Достовірність повідомлення $x \in X$ означає, що $x \in X^+$.

Твердження 5.

Всі стани $\omega' \in \Omega(x)$ рівноправні як «претенденти» на істинний стан ω^*, тобто ймовірність $P\{\omega' = \omega^* \mid \omega' \in \Omega(x)\} = |\Omega(x)|^{-1}$, $x \in X^+$. Тут і надалі $|A|$ означає потужність множини A. Припускається, що користувач обізнаний про відомості, зміст яких надано в *Твердженні 4* та *Твердженні 5*. Іншими словами, користувач знає, що спостерігач направляє йому достовірні повідомлення і всі вказані в ньому стани із множини $\Omega(x)$ рівноправні як претенденти на істинний стан об'єкта спостереження.

Твердження 6.

До отримання повідомлення $x \in X$ від спостерігача користувач володіє апріорною інформацією про стан об'єкта спостереження, яку можна представити в вигляді наявності в нього апріорного повідомлення $x_0 \in X$.

Твердження 7.

Повідомлення $x_0 \in X$ достовірне ($\omega^* \in \Omega(x_0)$) і всі стани $\omega' \in \Omega(x_0)$ рівноправні як претенденти на ω^*, тобто

$$P\{\omega' = \omega^* \mid \omega' \in \Omega(x_0)\} = |\Omega(x_0)|^{-1}.$$

Цінність інформації визначається користувачем. Це робиться наступним чином. До отримання повідомлення $x \in X$ від спостерігача користувач випадково і рівноймовірно отримує $\omega^{'} \in \Omega(x_0)$ і діє на основі припущення про те, що $\omega^{'} = \omega^*$. Зауважимо, що відповідно до *Твердження 7* ця стратегія дій користовуча є цілком обґрунтованою. З урахуванням наведених вище припущень його середній «дохід» визначається наступним чином:

$$V(x_0) = \frac{\alpha}{|\Omega(x_0)|} + (1 - \frac{1}{|\Omega(x_0)|})\beta.$$

При отриманні $x \in X^+$ користувач, знаючи що $\omega^* \in \Omega(x) \cap \Omega(x_0)$, випадково і рівноймовірно обирає $\omega^{'} \in \Omega(x) \cap \Omega(x_0)$ і отримує наступний середній дохід:

$$V(x \mid x_0) = \frac{\alpha}{|\Omega(x) \cap \Omega(x_0)|} + (1 - \frac{1}{|\Omega(x) \cap \Omega(x_0)|})\beta.$$

Під цінністю інформації, яка знаходиться в повідомленні $x \in X^+$, при наявності апріорних відомостей $x_0 \in X$ мається на увазі різниця:

$$S(x \mid x_0) = V(x \mid x_0) - V(x_0).$$

Таким чином маємо наступне співвідношення:

$$S(x \mid x_0) = C_{\alpha\beta}(\frac{1}{|\Omega(x) \cap \Omega(x_0)|} - \frac{1}{|\Omega(x_0)|}),$$

$$\text{де } C_{\alpha\beta} = \alpha - \beta > 0$$

Зауважимо, що цінність інформації знаходиться в наступних межах:

$$0 \leqslant S(x \mid x_0) \leqslant C_{\alpha\beta}(1 - \frac{1}{|\Omega(x_0)|}),$$

причому $S(x \mid x_0) = 0$, тоді і тільки тоді, коли $\Omega(x_0) \subseteq \Omega(x)$, і $S(x \mid x_0) = C_{\alpha\beta}(1 - \frac{1}{|\Omega(x_0)|})$, тоді і тільки тоді, коли $|\Omega(x) \cap \Omega(x_0)| = 1$.

1.5 Застосування теорії цінності інформації в криптографії

Теорія цінності інформації в криптографії є дуже важливим інструментом для розуміння того, як захищати інформацію від

несанкціонованого доступу. Згідно з теорією цінності інформації, інформація має вартість і цю вартість можна визначити на основі її корисності для потенційних користувачів.

У криптографії теорія цінності інформації застосовується для визначення рівня захисту інформації. За допомогою теорії цінності інформації можна визначити, яка інформація є найбільш цінною, і тому потребує найбільшого рівня захисту.

Криптографічні алгоритми використовуються для захисту інформації від несанкціонованого доступу. Застосування теорії цінності інформації дозволяє визначити, який рівень захисту потрібно застосовувати для конкретної інформації. Наприклад, для конфіденційної інформації, такої як особисті дані або фінансові дані, необхідно використовувати найбільш потужні криптографічні алгоритми, щоб захистити їх від несанкціонованого доступу.

Теорія цінності інформації в криптографії також допомагає визначити, який рівень ризику несанкціонованого доступу може бути прийнятним для конкретної інформації.

Крім того, теорія цінності інформації використовується для визначення ефективності криптографічних алгоритмів. Чим більше вартість інформації, тим складніше має бути розшифрування криптограми, тому що використовувані алгоритми повинні бути досить складними, щоб забезпечити належний рівень захисту. Таким чином, можна використовувати теорію цінності інформації для оцінки ефективності криптографічних алгоритмів і для розробки нових, більш потужних алгоритмів, щоб захистити найбільш цінну інформацію.

Крім того, теорія цінності інформації також використовується для вирішення проблеми цілісності інформації. Цілісність інформації означає, що дані не були змінені неправомірно, або що вони не були підмінені. Це може бути досягнуто за допомогою цифрових підписів, які базуються на криптографічних алгоритмах. Цифровий підпис може бути використаний для перевірки, що дані були підписані відповідною особою або організацією,

і що дані не були змінені після підписання.

Отже, теорія цінності інформації грає важливу роль у криптографії, оскільки допомагає забезпечити конфіденційність, цілісність та доступність інформації. Використання теорії цінності інформації дозволяє розробляти більш ефективні криптографічні протоколи і алгоритми, щоб захистити цінну інформацію від несанкціонованого доступу і зберегти її цілісність.

Прикладом застосування теорії цінності інформаії в криптографії є оцінка ефективності криптографічних систем за допомогою теорії цінності інформації.

Наведемо деякі теоретичні відомості із криптографії

Нехай шифр - це трьохкомпонентна універсальна алгебра:

$$A =< X, K, Y; f >,$$

де X – це скінчена множина відкритих текстів, K - це скінчена множина ключів, Y - скінчена множина шифротекстів; f – функція шифрування, відображення декартового добутку $X \times K$ на Y, тобто $f : X \times K \to Y$.

При зашифруванні повідомлення $x \in X$ обирається ключ $k \in K$ та визначається шифротекст $y = f(x, k)$. У строни, що отимує повідомлення, ключ k відомий. Отримавши шифрований текст $y \in Y$, користувач розшифровує його виористовуючи обернену функції шифрування:

$$x = f^{-1}(y, k).$$

Для однозначного розшифрування вимагається наступна умова: відображення $f : X \times K \to Y$ повинне бути ін'єктивним при будь-якому $k \in K$. Звідси випливає нерівність: $|X| \leq |Y|$.

Супротивник, перехопивши шифротекст $y \in Y$, намагається його дешифрувати, тобто визначити початковий відкритий текст $x \in X$. Відносно супротивника зробимо такі припущення:

1) Супротивник володіє інформацією щодо шифру A.

2) Супротивник не володіє інформацією щодо ключа, котрий використовувався під час шифрування.

3) Супротивник володіє необмеженими обчислювальними потужностями.

Останнє припущення відповідає методологічним положенням К. Шеннона у його дослідженні питання теоретичної стійкості шифрів. Воно значить, що при дешифруванні супротивник може використати метод повного перебору по усім ключам множини K.

Шифр A називається (X, Y)-транзитивним, якщо для будь-яких $x \in X$, $y \in Y$ існує такий ключ $k \in K$, що $y = f(x, k)$. Тобто, якщо шифр A є (X, Y)-транзитивним, то для будь-якого перехопленого $y \in Y$ маємо: $X(y) = X$. Шифри, які задовільнять такій властивості можна назвати «досконалими». Вони існують, але мають дуже обмежене застосування. Зауважимо, що для «досконалих» шифрів: $|X| \leq |Y| \leq |K|$.

Досконалий шифр, у якому $|X| = |Y| = |K|$ називається мінімальним шифром. Нехай $|x|$ – довжина повідомлення x у деякому алфавіті,$|k|$ – довжина ключа. Тоді у досконалому шифрі $|k| > |x|$, а у мінімальному досконалому шифрі $|k| \geq |x|$.

У інформаційно-аналітичній системі, що була описана вище, має місце наступне:

$$|X| = 2^{|\Omega|} - 1.$$

Відповідно, у мінімальному досконалому шифрі $|K| = 2^{|\Omega|} - 1$. У подальшому будемо виходити з припущення, що $2^{|\Omega|} \gg 1$. Тоді довжина ключа приблизно складає $|k|_2 = |\Omega| = log_2|K|$.

Введемо в структуру захисту інформації цілісні міркування. Супротивник, що перехопив повідомлення $y \in Y$, визначає множину

$$\Omega(y) = \bigcup_{x \in X(y)} \Omega(x).$$

Він знає, що $\omega^* \in \Omega(y)$. Розглянемо як саме визначається цінність перехвату повідомлення $y \in Y$. Прагнучі нанести користувачу максимальної шкоди, спостерігач випадково обирає $\omega \in \Omega(y)$ та діє з

припущенням, що $\omega^* = \omega$. У цьому випадку для спостерігача цінністю перехопленого шифрованого тексту є

$$S(y) = c_{\alpha\beta}(\tfrac{1}{|\Omega(y)|} - \tfrac{1}{|\Omega|}).$$

Припуская, що апріорно спостерігач мав відомості $x_0 = E$; α - прибуток спостерігача у випадку успіху ($\omega^* = \omega$) у нанесенні шкоди користувачу; β – втрати при невдачі ($\omega^* \neq \omega$), тобто величини α, β визначаються не користувачем, а спостерігачем.

Твердження.

Шифр A є S-досконалим тоді і тільки тоді, коли для будь-якого $y \in Y : S(y) = 0$.

Наслідок.

Якщо шифр A є досконалим, то він є і S-досконалим.

Твердження.

Шифр A спряжено-(X, Y)-транзитивним, якщо для будь-якого $y \in Y$ виконується умова: якщо $x \in X(y)$, то $\neg x \in X(y)$, де $\neg x$ – спряжене до x повідомлення, тобто $\Omega(\neg x) = \Omega \setminus \Omega(x)$.

Наслідок.

Якщо шифр A є спряжено-(X, Y)-транзитивним, то він є і S-досконалим.

Розглянемо ще одну можливість використання поняття цінності інформації при дослідженні питання криптографічної стійкості шифрів. При цьому будемо спиратись на ймовірнісну модель шифру

$$B =< A; P(X), P(K) >,$$

де $P(X), P(K)$ – розподіли ймовірностей на множинах X, K та відповідно:

$$P(X) = P(x), x \in X,\ P(K) = P(k), k \in K.$$

Відкриті тексти та ключі – незалежні, тобто $P(x, k) = P(x) \cdot P(k)$, $x \in X$, $k \in K$. Позначимо $K(x, y) = k \in K \mid f(x, k) = y$. Тоді розподіли P(X) та $P(K)$ однозначно індукують розподіли $P(Y) = P(y), y \in Y$:

$$P(Y) = \sum_{x \in X} P(x) \cdot \sum_{k \in K(x,y)} P(k).$$

Розглянемо «практичну стійкість шифрів». При цьому дослідимо проблему практичної стійкості так званих «вільних шифрів» (або «програмних шифрів»), які в наші дні в основному і використовуються. Ці шифри характеризуються рівністю:

$$|Y| = |X| \cdot |K|.$$

Таким чином, за будь-яким $y \in Y$ «ховається» єдина пара (x, k), $x \in X$, $k \in K$, яка і могла привести до появи $y = f(x, k)$.

Звідси випливає, що існує принципова можливість однозначного визначення $x \in X$, який «ховається» за перехватом $y \in Y$.

Враховуючи все вище описане, вільний шифр являється повною протилежністю досконалим шифрам. Дійсно, при використанні вільного шифру маємо: $\Omega(y) = \Omega(x)$, де $y = f(x, k)$.

Далі, цінність перехвату $y \in Y$ дорівнює цінності відповідного відкритого повідомлення. Такі шифри мають практичне значння тому, що сама процедура визначення $x \in X$ по $y \in Y$ може займатиьдуже великий час. За цей час інформація $x \in X$ знецінюється в результаті «старіння».

Термін «практична стійкість» шифра А як раз і має той сенс, що процедура визначення $x \in X$ по $y \in Y$ вимагає настільки великих витрат часу, що практично ці затрати становляться невиправданими.

Висновки до розділу 1

В данному розділі наведено необхідні теоретичні матеріали щодо теорії цінності інормації, моделей цінності інормації, моделі цінності інформації за Г. П. Шанкіном та застосування теорії цінності інформації в криптографії. Проведено огляд опублікованих джерел по темі дослідження.

2 УЗАГАЛЬНЕНА МОДЕЛЬ ЦІННОСТІ ІНФОРМАЦІЇ НА ШИРОКИЙ СПЕКТР РОЗПОДІЛІВ В СИСТЕМІ ПЕРЕДАЧІ ІНФОРМАЦІЇ В УМОВАХ НЕВИЗНАЧЕНОСТІ

В цьому розділі запропонована модель цінності інформаційної системи з передачею інформації по каналу зв'язку в умовах невизначеності, в якій розподіли на множині станів об'єкта належать широкому класу розподілів, що містить також рівноймовірний розподіл. Ця модель узагальнює інформаційно-аналітичну систему з визначенням цінності інформації, розроблену в роботі Г. П. Шанкіна. При нерівноймовірному початковому розподілі станів об'єкту можна створити більш оптимальні стратегії поведінки користувача, які збільшують ймовірність вгадування істинного стану і збільшують дохід користувача.

2.1 Узагальнена модель передачі інформації в умовах невизначеності

Розглядається інформаційно-аналітична система. В цій системі присутні:

1) Об'єкт спостереження, який може знаходитися в одному з n станів з множини усіх можливих станів $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, ω_* - істинний стан об'єкта спостереження.

2) $P = \{p_1, p_2, ..., p_n\}$, $P(\omega_i) = p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$ - це ймовірнісний розподіл, заданий апріорно на множині станів Ω.

3) $X = X^+ \cup X^-$ – множина всіх можливих повідомлень спостерігача.

4) X^+ - множина достовірних повідомлень $x \in X^+$, для яких $\omega \in \Omega(x)$

5) X^- - множина недостовірних повідомлень $x \in X^-$, для яких $\omega \notin \Omega(x)$

6) Спостерігач, який має інформацію (повідомленя) x, змістом якого є інформація про підмножину станів $\Omega \subseteq \Omega$, яка містить або не містить істинний стан.

7) Користувач, який використовує отриману від спостерігача інформацію (апостеріорну інформацію) і свою інформацію про істинний стан об'єкта спостереження (апріорна інформація) для вирішення певної задачі.

8) Канал зв'язку, по якому спостерігач передає повідомлення користувачу.

9) Криптосистема шифрування повідомлень.

10) Супротивник, який перехоплює повідомлення і хоче отримати доступ до інформацію, або передати користувачу дезінформацію.

Будемо вважати, що успіх користувача у вирішенні задачі залежить від інформації про ω^* - істинний стан об'єкта спостереження. Успіх у вирішенні користувачем задачі будемо вимірювати (оцінювати) середнім доходом користувача за такими припущеннями:

- якщо користувач при вирішенні задачі використовує інформацію про істинний стан ω^* об'єкта спостереження, тоді корисувач отримує отримує дохід $\alpha > 0$,

- якщо користувач використовує стан $\omega \neq \omega^*$, тоді користувач отримує дохід $0 \leq \beta < \alpha$.

Розподіл $P = \{p_1, p_2, ..., p_n\}$ індукує розподіли на множині станів $\Omega(x) \subseteq \Omega$ для кожного повідомлення x, а саме $\forall \omega_i \in \Omega(x)$ умовна ймовірність дорівнює

$$P(\omega_i \mid x) = P(\omega_i \mid \Omega(x)) = \frac{p_i}{P(\Omega(x))}.$$

Позначимо цю ймовірність як

$$p_i(\Omega(x)) \text{ або } p_i(x) = P(\omega_i \mid x).$$

Для довільного повідомлення $x \in X$ позначимо

$$p^*(x) = max_{\omega_i \in \Omega(x)} p_i(x) \text{ і } p^* = max_{\omega \in \Omega} p_i = max_{i=1,...,n} p_i.$$

Буудемо вважати змістом декількох повідомлень $x_1, x_2, ..., x_r$ повідомлення $(x_1, x_2, ..., x_r)$, зміст якого є перетином змістів окремих повідомлень $\Omega(x_1, x_2, ..., x_r) = \cap_{i=1}^{r} \Omega(x_i)$.

Припущення щодо розподілу $P = \{p_1, p_2, ..., p_n\}$ на множині станів Ω:

1) Якщо $\Omega(x_2) \subseteq \Omega(x_1)$, $p^*(x_2) \geq p^*(x_1) \geq p^*$.

2) Якщо має місце строге включення $\Omega(x_2) \subset \Omega(x_1) \subset \Omega$, то $p^*(x_2) > p^*(x_1) > p^*$.

Приклади існування розподілів для другого припущення:

1) Очевидним прикладом існування розподілу, що задовольняє другому припущеню є рівноймовірний розподіл

$$P = \{p_1, p_2, ..., p_n\},\ P(\omega_i) = \frac{1}{n},\ \sum_{i=1}^{n} p_i = 1.$$

2) Побудуємо розподіл $P = \{p_1, p_2, ..., p_n\}$, в якому ймовірності лінійно залежать від індексу

$$p_{i+1} - p_i = \Delta = const > 0,\ i = 1, 2, ..., n-1.$$

Запишемо ймовірність у вигляді $p_i = \frac{(k+i)\Delta}{q}$, $i = 1, ..., n$, $k = 0, 1, 2, ...$ - будь-який фіксований параметр. Звідки знаходимо

$$p_i = \frac{2(k+i)}{n(n+2k+1)},,\ i = 1, ..., n,\ k = 0, 1, 2, ...,$$

$$\Delta = p_(i+1) - p_i = \frac{2}{n(n+2k+1)},\ i = 1, 2, ..., n-1.$$

Звідси можно побачити, що умови другого припущення виконуються, наприклад при $k = n = 5$ маємо:

$$p_1 = \frac{6}{40},\ p_2 = \frac{7}{40},\ p_3 = \frac{8}{40},\ p_4 = \frac{9}{40},\ p_5 = \frac{10}{40}$$

Зауваження 1.

Очевидно, якщо умови другого припущення для деякого розподілу із другого прикладу виконуються, то умови будуть виконуватися при будь-якій перестановці індексів у ймовірностях розподілу із другого прикладу.

Зауваження 2.

При $k \to \infty$ розподіл із другого прикладу збігається до рівноймовірного розподілу із першого прикладу. Існує багато інших

ймовірнісних розподілів, для яких виконуються умови другого припущення.

2.2 Стратегія поведінки користувача і спостерігача в узагальненій моделі цінності інформації

1) Користувач має деяку апріорну достовірну інформацію x_0 щодо стану об'єкту спостереження зі змістом $\Omega(x_0)$, $\omega^* \in \Omega(x_0)$, від якого залежить успіх у виконанні певної задачі користувачем і досягнення цілі.

2) Спостерігач посилає користувачу одно або декілька достовірних повідомлень $x_1, x_2, ..., x_r \in X^+$ і має на меті сприяти успіху користувача у вирішені певної задачі. Будемо вважати, що в достовірних повідомленнях стани відмінні від істинного вибираються випадково серед інших станів і незалежно від вибору в інших подомленнях.

3) Якщо користувач не отримував повідомлень від спостерігача, то для виконання задачі він діє, вважаючи, що об'єкт спостереження має істинний стан ω^*, який відповідає $p^*(x_0)$, тобто стан з максимальною ймовірністю при умові $\Omega(x_0)$. Якщо таких станів декілька, то кандидат у істинний стан вибирає випадковим чином серед них. При такій стратегії користувач максимізує свій середній дохід з інформацією x_0.
Дохід буде дорівнювати:

$$V(x_0) = \alpha p^*(x_0) + \beta(1 - p^*(x_0)).$$

4) Користувач довіряє інформації, отриманої від спостеріача, тобто він на сто відсотків впевнений, що зміст отриманого повідомлення включає в себе істинний стан. Якщо користувач отримав повідомлення x від спостерігача, то для виконання задачі він вибирає істинний стан ω^* об'єкта спостереження, який відповідає $p^*(x, x_0)$, тобто стан з максимальною ймовірністю при умові $\Omega(x, x_0)$. При такій стратегії користвач ефективно використає початковий розподіл на станах, свою апріорну інформацію і інформацію в повідомленні спостерігача та

отримає середній дохід:

$$V(x, x_0) = \alpha p^*(x, x_0) + \beta(1 - p^*(x, x_0)).$$

5) З використанням інформації від спостерігача користувач має збільшення доходу на величину:

$$V(x, x_0) - V(x_0) = (\alpha - \beta)(p^*(x, x_0) - p^*(x_0)).$$

Будемо вимірювати цінність інформації саме величиною збільшення отриманого доходу при її використанні. Позначимо різницю $\alpha - \beta = C_{\alpha\beta}$.

Введемо означення цінності інформації повідомлення x. Цінністю інформації, що міститься в повідомленні x, із урахуванням нерівноймовірного розподілу ймовірностей на множині станів Ω, для користувача і спостерігача, при наявності у користувача апріорної інформації x_0, називається величина:

$$S(x \mid x_0) = V(x, x_0) - V(x_0) = C_{\alpha\beta}(p^*(x, x_0) - p^*(x_0)).$$

Запропонована модель визначення цінності інформації з новою стратегією дій користувача узагальнює інформаційно-аналітичну модель Г. П. Шанкіна на широкий клас ймовірнісних розподілів на множині станів об'єкту спостереження і дає можливість збільшити дохід користувача. Якщо ймовірнісний розподіл на множині станів рівноймовірний, то висновки за узагальненою моделлю співпадають з відповідними висновками моделі Шанкіна.

2.3 Властивості цінності інформації в узагальненій моделі із використанням нерівноймовірного розподілу на множині усіх станів

1) Цінність інформації в узагальненій моделі задовольняє нерівностям

$$0 \leq S(x \mid x_0) \leq C_{\alpha\beta}(1 - p^*(x_0)).$$

Дійсно, з умов другого припущення про розподіли випливає, що $0 \leq S(x \mid x_0)$. При цьому $S(x \mid x_0) = 0$ тоді і тільки тоді, коли $\Omega(x_0) \subseteq \Omega(x)$ і користувач не отримує додаткової інформації про істинний стан. Цінність інформації максимальна, коли вона дозволяє однозначно визначити істинний стан, тобто коли $|\Omega(x_0) \cap \Omega(x)| = 1$ і $p^*(x, x_0) = 1$.

2) Будемо позначати достовірне повідомлення $x \in X^+$, у якого $|\Omega(x)| = 1$, як і в звичайній моделі через e, а $x \in X^+$ зі змістом $\Omega(x) = \Omega$ як E. Тоді виконується:

$$\forall x \in X^+\colon S(e \mid x) = C_{\alpha\beta}(1 - p^*(x_0)),\ S(x \mid e) = 0,$$

$$\forall x \in X^+\colon S(E \mid x) = 0,\ S(x \mid E) = C_{\alpha\beta}(p^*(x) - p^*).$$

Позначимо $S(x \mid E) = S(x)$ і будемо називати безусловною цінністю інформації.

3) В узагальненій моделі, по аналогії зі звичайною моделлю інформаційно-аналітичної системи, виводиться формула

$$S(x_1, x_2, ..., x_r) =$$
$$= S(x_1) + S(x_2 \mid x_1) + S(x_3 \mid x_1, x_2) + ... + S(x_r \mid x_1, x_2, ..., x_{r-1}).$$

В узагальненій схемі будемо називати повідомлення $x_1, x_2 \in X^+$ рівноцінними, якщо $p^*(x_1) = p^*(x_2)$. Чим менші ці ймовірності, тим менше точність відповідного повідомлення.

4) Поняття цінності інформації $S^-(x)$, якої не вистачає повідомленню $x \in X^+$ для однозначного визначення істинного стану $\omega^* \in \Omega$, переноситься і на узагальнену модель із нерівноймовірним розподілом на станах Ω. Повідомлення $x^{'} \in X^+$, яке містить таку інформацію, має властивість $|\Omega(x) \cap \Omega(x^{'})| = 1$. Тоді знаходимо

$$S^-(x) = S(x, x^{'}) - S(x) = C_{\alpha\beta}(1 - p^*) - C_{\alpha\beta}(p^*(x) - p^*) = c_{\alpha\beta}(1 - p^*(x)),$$

$$S^-(x, x^{'}) = C_{\alpha\beta}(1 - C_{\alpha\beta}(1 - p^*(x, x^{'}))),$$
$$S(x \mid x^{'}) = C_{\alpha\beta}(p^*(x, x^{'}) - p^*(x)).$$

Склавши дві останні формули можна отримати наступне твердження. Для будь-яких $x, x^{'} \in X^+$ в узагальненій моделі виконується співвідношення

$$S^-(x) = S^-(x, x') + S(x' \mid x).$$

В звичайній моделі це співвідношення, введене для рівнойморівного випадку, називається «законом збереження інформації».

2.4 Дезінформація в узагальненій моделі цінності інформації

1) В узагаленій моделі користувач довіряє спостерігачу і проводить при розв'язанні своєї задачі страткгію поведінки, яку описано вище.

2) Спостерігач виступає у ролі супротивника і намагається своїми повідомленнями нанести максимальних втрат користувачу.

Реалізація стратегії супротивника для дезінформації користувача

Мета супротивника – нанести користувачу максимальні середні втрати з використанням фальшивих повідомлень. Розглянемо по аналогії зі звичайною моделлю більш простий для аналізу випадок: супротивник знає апріорну інформацію користувача і використовує також її для створення дезінформації. В залежності від змісту достовірної інформації в повідомленнях $x_0, x_c \in X^+$ відповідно у користувача і спостерігача стратегією спостерігача в побудові хибного повідомлення $x^- \in X^-$ для користувача можна розбити на 2 таких випадки.

1) Виконується $\Omega(x_0) \not\subseteq \Omega(x_c)$. Оптимальною стратегією супротивника в цьому випадку буде вибір повідомлення $x^- \in X^-$, яке задовольняє умовам:

$$\Omega(x_0) \cap \Omega(x^-) = \varnothing \text{ та } \Omega(x^-) \subseteq \Omega \setminus (\Omega(x_0) \cap \Omega(x_c))$$

Завдяки першій умові користувач не розпізнає обману з боку супротивника. При $\Omega(x_0) \cap \Omega(x^-) = \varnothing$ користувач одразу побачив би, що надіслане повідомлення $x^- \in X^-$ недостовірне.

За другою умовою істинний стан $\omega^* \notin \Omega(x_0) \cap \Omega(x^-)$ і при виконанні задачі користувач користується інформацією про хибний стан об'єкта спостереження та понесе такі втрати:

$$V(x_0) - V(x_0, x^-) = \alpha p^*(x_0) + \beta(1 - p^*(x_0)) - \beta =$$
$$= (\alpha - \beta)p^*(x_0) = c_{\alpha\beta}p^*(x_0).$$

2) Виконується $\Omega(x_0) \subseteq \Omega(x_c)$. Це означає, що супротивник не має більш точної інформації про істинний стан об'єкта спостереження ніж користувач і не зможе сформувати повідомлення з дезінформацією, із-за якого користувач понесе втрати.

Висновки до розділу 2

В розділі запропоновано та формально описано узагальнення моделі цінності інформації за Г. П. Шанкіном на широкий спектр розподілів в системі передачі інформації в умовах невизначеності. Наведено теоретичний опис цієї моделі. При використанні нерівноймовірних розподілів запропонована більш оптимальна стратегія поведінки користувача, що в свою чергу максимізую його дохід та цінність інформації.

3 РЕАЛІЗАЦІЯ УЗАГАЛЬНЕНОЇ МОДЕЛІ ЦІННОСТІ ІНФОРМАЦІЇ ІЗ РІЗНИМИ ВХІДНИМИ ДАНИМИ ТА ДОСЛІДЖЕННЯ РЕЗУЛЬТАТІВ ЇЇ РОБОТИ

В данному розділі наведена реалізація чотирьох варіантів узагальненої моделі цінності інформації, опис якої наведено в попередньому розділі. Наведено опис функціоналу, створеного веб-застосунку. Також проведено аналіз отриманих результатів.

3.1 Загальний опис чотирьох варіантів узагальненої моделі цінності інформації

Для програмної реалізації обрана узагальнена модель цінності інформації, яка описана в минулому розділі. Запропоновано чотирb різні варінти моделі, які відрізняються вхідними даними та поведінкою користувача.

Загальна схема роботи програмної реалізfції наступна:

1) Будується множина станів із заданим розподілом ймовірностей на цій множині, тобто кожен стан має свою ймовірність бути істинним.

2) Згідно із заданим розподілом обирається та фіксується стан в якому знаходиться об'єкт спостереження, тобто істиний стан.

3) Випадковим чином із множини усіх станів створюються повідомлення користувача та повідомлення спостерігача, які включають в себе попередньо зафіксований істиний стан. Знаходиться перетин цих повідомлень.

4) Проводяться декілька раундів вгадування істинного стана на повідомленні користувача та перетені повідомлення користувача із повідомленням спостерігача. Отримані результати вгадувань являють собою ймовірність вгадати істиний стан, маючи апріору інформацію та

апріорну і апостеріорну, відповідно.

5) Здійснюється обчислення доходів для апріорної та апостеріорної інформацій.

6) Віднявши доходи апріорної інформації від апріорної та апострeріорної, отримуємо цінність інформації, яку відправив спостерігач користувачу.

7) Повторюємо етапи 2–6 декілька разів для того, щоб отримати дані для подальшого усереднення.

Наведена схема актуальна для всіх чотирьох реалізованих варіантів узагальненої моделі цінності інформації.

Наведемо опис чотирьох реалізованих варіантів узагальненої моделі цінності інформації:

Варіант 1.

Перша модель являє собою частковий випадок узагальненої моделі, коли розподіл ймовірностей на всіх станах об'єкта спостереження Ω є рівноймовірним, тобто параметр розподілу $k \to \infty$. В такому випадку оптимальною поведінкою користувача буде випадковий та рівноймовірний вибір стану із повідомлення в якості кандидата на істинний стан. При такій стратегії, із урахуванням рівноймовірності на станах об'єкта спостереження, поведінка кристувача є оптимальною, тобто він максимізує свій середній дохід.

Варіант 2.

Другий варінт моделі відрізняється від першого тим, що тут задано нерівноймовірний розподіл на множині станів Ω. При цьому істиний стан також обирається згідно цього розподілу, тобто стан із більшою ймовірністю має більше шансів стати істинним. Як і в першому варінті, повіденкою користувача є випадковий та рівноймовірний вибір кандидата на істинний стан із повідомлення, без урахування заданого розпділу.

Варіант 3.

Третій варіант узагальненої моделі цінності інформації передбачає нерівноймовірний розподіл на множині станів Ω. Але при цьому повіденка

користувача відрізнється від першого та другого варіантів, а саме: в якості кандидата на істинний стан обирається стан із найбільшою ймовірністю в повідомленні.

Варіант 4.

Четвертий та останній варіант моделі являє собою комбінацію другого та третього варіантів. В цьому варінті також задан нерівноймовірний розподіл на множині станів Ω. При цьому, поведінка користувача дещо різниться. Знаючи розподіл на усіх станах, та отримавши пвідомлення, користувач обирає підмножину найбільш ймовірних станів із повідомлення, та серед цієї підмножини вже обирає кандидата на істиний стан об'єкта спостереження.

Зауваження для варіантів 3-4.

У варінтах моделі 2, 3 та 4 використовується нерівноймовірний розподіл на множині станів Ω, який формально описується наступним чином:

$$P = \{p_1, p_2, ..., p_n\},\ p_i = \frac{2(k+i)}{n(n+2k+1)},\ i = 1, ..., n,\ k = 0, 1, 2, ...,\ \sum_{i=1}^{n} p_i = 1.$$

При цьому обирати параметр розподілу k можно будь-яким, отримуючу різні розподіли.

3.2 Опис функціоналу створеного веб-застосунку

Під час програмної реалізації створено веб-застосунок, який дозволяю зручним чином змінювати вхідні параметри моделі та поведінку користувача, обираючи різні варіанти моделі. Веб-застосунок складається із трьох функціональних модулей. Ознайомитися із програмною реалізацією можна за посиланням: Information Value Model.

Перший модуль

Перший модуль надає можливість обирати такі важливі параметри як вид розподілу та повіденінку користувача, тото перемикатися між різними варіантами роботи моделі, які описані вище. Цей модуль

складається із секції вибору варінта моделі та секції для завдання параметра розподілу k.

За замовчуванням задані такі параметри: варіант моделі - 1, параметр розподілу $k = 0$.

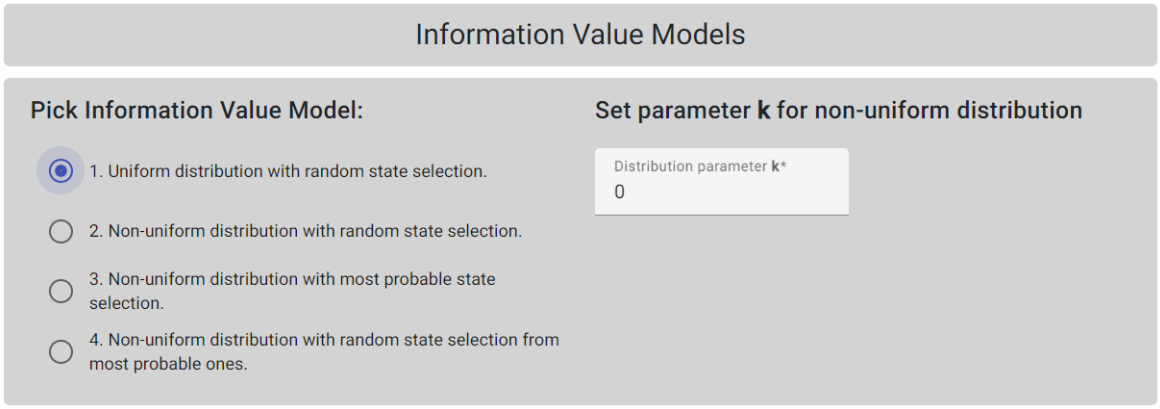

Рисунок 3.1 – Перший модуль веб-застосунку.

Другий модуль

Другий функціональний модуль дозволяє змінювати основні параметри моделі цінності інформації. Цей модуль складається із декількох полей для ввода чисел, які дозволяють керувати наступними параметрами моделі:

1) Параметр α - ефективність дій користувача у випадку, коли йому вдалося вгадати істинний стан об'єкта спостереження.

2) Параметр β - ефективність дій користувача у випадку, коли йому не вдалося вгадати істинний стан об'єкта спостереження.

3) Кількість експерементів - скільки разів буде проводиться експеремент, тобто обчислюватися цінність деякого повідомлення, для подальшого усереднення по цій кількості.

4) Кількість «вгадувань» - скільки разів буде проводитися вгадування зафіксованого істинного стану під час одного експеременту для отримання ймовірносей вгадування і подальшого обчислення доходів.

5) Кількість усіх станів об'єкта спостереження Ω.

6) Розмір повідомлення користувача (апріорна інформація), тобто

кількість станів, які містяться в цьому повідомлені.

7) Розмір повідомлення спостерігача (апостеріорна інформація), тобто кількість станів, які містяться в цьому повідомлені.

8) Відсоток повідомлення, який буде обиратися для варіанта моделі під номером чотири, тбто яка кількість найбільш ймовірних станів із повідомлення увійде до підмножини.

Також в цьому модулі присутній перемикач, який дозволяє обирати довжину та ноповнення повідомлень випадковим чином. При цьому поля для ввода довжин повідомлень користувача та спостерігача замінюються на поля для ввода порогів для входження стану у повідомлення користувача та спостерігача відповідно.

При використанні режиму моделі із випадковими довжинами повідомлень, Обирається поріг входження стану в повідомлення $0 \leq T \leq 1$. Для кожного стану $\omega \in \Omega$ генерується випадкове число $0 \leq t \leq 1$. Якщо $t \leq T$, то стан додається в повідомлення. Це ще один приклад узагальнення моделі Г. П. Шанкіна, так як при використанні такої стратегії будови повідомлень забезпеується максимальні випадковість та рівноймовірність отриманих повідомлень.

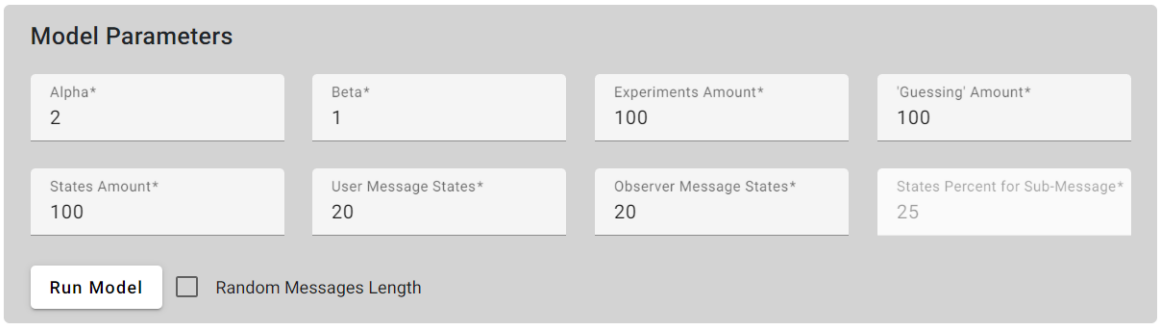

Рисунок 3.2 – Другий модуль веб-застосунку із фіксованими довжинами повідомлень.

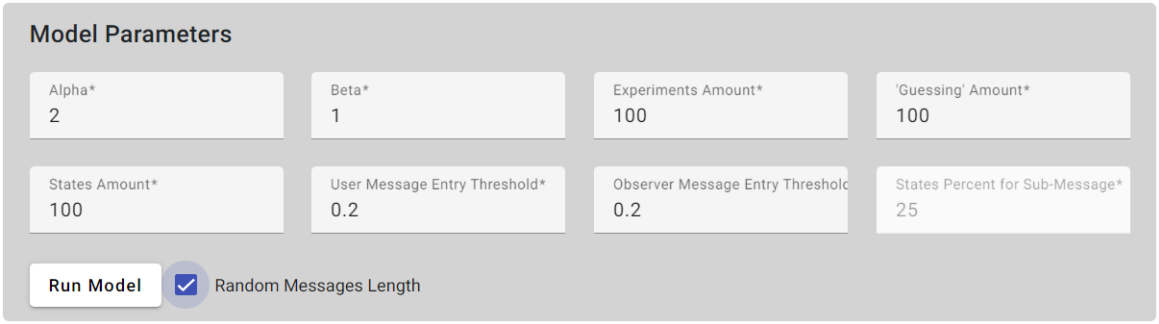

Рисунок 3.3 – Другий модуль веб-застосунку із випадковими довжинами повідомлень.

Третій модуль

Третій модуль являє собою таблицю, в яку виводяться результати роботи моделі. Таблиця складається із шесте стовпців: номер моделі, середній розмір повідомлення користувача, середній розмір повідомлення спостерігача, середній апріорний дохід користувача, середній апостеріорний дохід користувача, середня цінність інформація. Всі ці величини усереднюються по заданій кількості експерементів.

3.3 Аналіз отриманих результатів

Під час дослідження узагальненої моделі цінності інформації проведенно декілька раундів експереметів чотирьох варіантів моделі. Експеременти проводилися із наступними параметрами: $\alpha = 11$, $\beta = 1$, кількість есперементів для усереднення - 100, кількість «вгадувань» під час експеременту - 100, загальна кількість станів об'єкта спостереження Ω - 100, фіксована довжина повідомлення користувача - 20, фіксована довжина повідомлення спостерігача - 20, поріг входження стану в повідомлення користувача - 20, поріг входження стану в повідомлення спостерігача - 20, процент найбільш ймовірних станів в підмножині для четвертого варіанта моделі - 50%.

Результати представлені у вигляді трьох таблиць. Для отримання результатів для кожної таблиці використовувались різні значення

параметра розподілу k.

Варіант Моделі	Середній апріорний дохід	Середній апостеріорний дохід	Середня цінність інформації
1	1.52	3.43	1.91
1	1.52	3.36	1.84
2	1.51	3.44	1.93
2	1.50	3.46	1.95
3	1.94	5.22	3.28
3	2.41	5.54	3.13
4	1.69	4.02	2.33
4	1.78	4.24	2.46

Таблиця 3.1 – Результати роботи моделі при заданому параметрі розподілу $k = 0$.

Варіант Моделі	Середній апріорний дохід	Середній апостеріорний дохід	Середня цінність інформації
1	1.50	3.45	1.95
1	1.50	3.48	1.98
2	1.50	3.49	1.99
2	1.50	3.45	1.95
3	1.82	4.51	2.68
3	2.12	4.81	2.70
4	1.64	4.03	2.40
4	1.62	4.03	2.41

Таблиця 3.2 – Результати роботи моделі при заданому параметрі розподілу $k = 20$.

Варіант Моделі	Середній апріорний дохід	Середній апостеріорний дохід	Середня цінність інформації
1	1.50	3.47	1.97
1	1.49	3.49	1.99
2	1.50	3.42	1.91
2	1.49	3.45	1.95
3	1.74	4.39	2.65
3	1.67	4.07	2.40
4	1.59	3.77	2.19
4	1.62	4.04	2.43

Таблиця 3.3 – Результати роботи моделі при заданому параметрі розподілу $k = 40$.

Як можна побачити третій варіант моделі дає найбільшу цінність інформації, а також і найбільші доходи від повідомлень. Звідси можна зробити наступний всновок: якщо розподіл ймовірностей на множині станів нерівноймовірний, то оптимальною поведінкою користувача буде вибір найбільш ймовірного стану в якості кандидата на істиний стан. Також слід зауважити, що із збільшенням параметру розподілу k результати роботи моделі «вирівнюються», що свідчить ро те, що розподіл ймовірностей стає більш схожим на рівноймовірний, і стратегія із вибором найбільш ймовірного стану не дає великої переваги.

Висновки до розділу 3

В розділі описана програмна реалізація узагальненої моделі ціннос інформації. Проведено аналіз отриманих результатів роботи моделі. Експерементально доведено оптимальність стратегії вибору найбільш ймовірного стану із повідомлень при наявності нерівноймовірного розділу на усіх станах об'єкта спостереження Ω.

ВИСНОВКИ

В роботі наведені теоретичні відомості щодо теорії цінності інформації та моделей цінності інормації. Розкриті можливі застосування теорії цінності інформації в криптографії. Після проведеного огляду теоретичних відомостей, із великого переліку запропонованих моделей цінності інформації обрана одна модель дл подальшого дослідження, а саме модель цінності інформації, запропонованая Г. П. Шанкіном. Після дослідження цієї моделі, запропоновано та теоретично описано узагальнення обранної моделі на широкий спектр розподілів в системі передачі інформації в умовах невизначеності. Також запропоновано декілька удоскональнь, як, наприклад, нова стратегія побудови повідомлень. В роботі також, окрім теоретичних викладок, зроблені експерементальні дослідження роботи узагальненної моделі цінності інформації. Створено веб-застосунок, який дозволяє зручним чином отримувати результати при різних вхідних параметрах, як, наприклад, поведінка користувача чи наявний розподіл на усіх станах об'єкта спостереження.

Таким чином, реалізовано чотири різні варіанта моделі, отримані та проаналізовані результати їх роботи. Як результат, експерементально доведено, що оптимальною поведінкою користувача, в умовах, якщо на усіх станах задано нерівноймовірний розподіл, буде вибір у якості кандидата на «істинний» стан об'єкта спостерження найбільш ймовірного стану із апріорної інформації, якщо мова йде про апріорний дохід, та із перетену апріорної інформації із повідомленням спостерігача, якщо мова йде про апостеріорний дохід. Таку стратегія користувача дає середній приріст цінності інформації у 50%. Якщо ж розподіл на станах об'єкта спостереження рівноймовірний, то оптимальною стратегією користувача залишається обирати випадковий стан. Звідси можна побачити, що узагальнення моделі цінності інформації на широкий спектр розподілів

дозволяє не тільки застосовувати цю модель на більш широкий перелік практичних задач, а ще й отримувати більший прирост доходу користувача.

Із кодом програмної реалізації можна ознайомитися за посиланням:
Shvyika GitHub, Diploma Master

Із веб-застосунком можна ознайомитися за посиланням:
Information Value Model

ПЕРЕЛІК ПОСИЛАНЬ

1. F. A. Hayek «The Use of Knowledge in Society», 1945 - 92 p.

2. Herbert Simon «Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization», 1957 - 123 p.

3. Морозевич А.Н. «Основи економічної інформатики», 1998 - 25 стр.

4. Романов В.П. «Інтелектуальні інформаційні системи в економіці», 2003 - 146 стр.

5. Кондаков Н.І. «Логічний словник – довідник», 1976 - 74 стр.

6. Богуш В.М., Кривуця В.Г., Кудін А.М. Інформаційна безпека. Термінологічний навчальний посібник, (2004) - 42 стр.

7. Архипов О.Є. «Критерії визначення можливої шкоди національній безпеці України у разі розголошення інформації, що охороняється державою»: моногр. // О.Є.Архипов, О.Є.Муратов. // – К.: Наук.-вид. відділ НА СБ України, 2011. – 195с.

8. Архипов О.Є. «Щодо методики ідентифікації та оцінювання активів системи інформаційних технологій» // О.Є. Архипов // Захист інформації. – №1(50), 2011. С. 42-47.

9. Архипов О.Є. «Застосування методології передбачення для оцінювання шкоди, заподіяної витоком секретної інформації» // О.Є.Архипов, І.П.Касперський // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – Вип.2(15). – К. 2007. – С.13-19.

10. Г. П. Шанкин «Теория ценности информации», 2004 г. - гл. 1, 6 і 7.