

ДОДАТОК А

Лістинг фрагментів коду програми

1) SiteDTO.cs

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Web;
namespace StrayMan.Models
{
public class SiteDTO
{
// идентификация страницы
public string SiteID { get; set; }
// адрес страницы в сети Интернет
public string Url { get; set; }
// Название документа
public string Title { get; set; }
// Описание документа
public string Description { get; set; }
// Тип документа
public string Type { get; set; }
// Ссылка Кэш документа
public string CacheLink { get; set; }
// SES документа
public float SES { get; set; }
}
}
```

2) AccountModels.cs

```
using System;
using System.Collections.Generic;
using System.ComponentModel.DataAnnotations;
using System.Globalization;
using System.Web.Mvc;
using System.Web.Security;
namespace StrayMan.Models
```

```

{
//изменение пароля
public class ChangePasswordModel
{
//текущий пароль
[Required]
[DataType(DataType.Password)]
[Display(Name = "Current password")]
public string OldPassword { get; set; }
//новый пароль
[Required]
[StringLength(100, ErrorMessage = "The {0} must be at least {2}
characters long.", MinimumLength = 6)]
[DataType(DataType.Password)]
[Display(Name = "New password")] 154
// Подтверждения нового пароля
[DataType(DataType.Password)]
[Display(Name = "Confirm new password")]
[Compare("NewPassword", ErrorMessage = "The new password and
confirmation password do not match.")]
public string ConfirmPassword { get; set; }
}
// для авторизации
public class LogOnModel
{
// имя пользователя
[Required]
[Display(Name = "User name")]
public string UserName { get; set; }
// пароль пользователя
[Required]
[DataType(DataType.Password)]
[Display(Name = "Password")]
public string Password { get; set; }
// запоминание имя и пароля
[Display(Name = "Remember me?")]

```

```

public bool RememberMe { get; set; }
}
// для регистрации
public class RegisterModel
{
// имя пользователя
[Required]
[Display(Name = "User name")]
public string UserName { get; set; }
// Email пользователя
[Required]
[DataType(DataType.EmailAddress)]
[Display(Name = "Email address")]
public string Email { get; set; }
// пароль пользователя
[Required]
[StringLength(100, ErrorMessage = "The {0} must be at least {2}
characters long.", MinimumLength = 6)]
[DataType(DataType.Password)]
[Display(Name = "Password")]
public string Password { get; set; }
// Подтверждения пароля
[DataType(DataType.Password)]
[Display(Name = "Confirm password")]
[Compare("Password", ErrorMessage = "The password and confirmation
password do not match.")]
public string ConfirmPassword { get; set; }
}
}

```

3) HomeController.cs

```

using System;
using System.Collections.Generic; 155

```

```
using System.Linq;
using System.Web;
using System.Web.Mvc;
using StrayMan.Models;
using StrayMan.DAL;
namespace StrayMan.Controllers
{
public class HomeController : Controller
{
// для доступа к данным
DataAccess _db = new DataAccess();
//для списка результатов
List<SiteDTO> results = new List<SiteDTO>();
public ActionResult Index()
{
return View();
}
//поиск результатов
public ActionResult Search(string query, int? page)
{
// проверка существования пользователя
if (User.Identity.IsAuthenticated)
{
//для получения результатов из прецедентов
results = _db.getResult(query, "htm");
ViewBag.query = query;
// для добавления прецедента (сохранение)
LocalPrecedentEntities _local = new LocalPrecedentEntities();
Precedent__case_ Pc = new Precedent__case_
{
CaseId = _local.Precedent__case_.Count() + 1,
Query = query,
Username = User.Identity.Name
};
_local.Precedent__case_.AddObject(Pc);
_local.SaveChanges();
}
```

```

}
else
{
// для получения результатов из индекса ИПС
results = _db.getResults(query, "htm");
ViewBag.query = query;
}
return View("Search", results);
}
public ActionResult About()
{
return View(); 156
}
}
4) DataAccess.cs

```

```

using System;
using System.Collections.Generic;
using System.Linq;
using System.Web;
using StrayMan.Models;
namespace StrayMan.DAL
{
public class DataAccess
{
ServerPrecedentManager.PrecedentManagerClient pmanager = new
ServerPrecedentManager.PrecedentManagerClient();
ServerAgentManager.IndexManagerClient imanager = new
ServerAgentManager.IndexManagerClient();
public List<SiteDTO> getResult(string query, string queryType)
{
// Получение результатов из прецедентов
var sResults = pmanager.getResult(query, queryType);
List<SiteDTO> results = new List<SiteDTO>();
results = (from r in sResults

```

```
select new SiteDTO
{
    CacheLink = r.CacheLink,
    Description = r.Description,
    SiteID = r.SiteID,
    Title = r.Title,
    Type = r.Type,
    Url = r.Url,
    SES = r.SES
}).ToList();
return results;
}

public List<SiteDTO> getResults(string query, string queryType)
{
    // Получение результатов из индекса ИПС
    var sResults = imanager.getResult(query, queryType);
    List<SiteDTO> results = new List<SiteDTO>();
    results = (from r in sResults
    select new SiteDTO
    {
        CacheLink = r.CacheLink, 157

        Description = r.Description,
```

ДОДАТОК Б
Слайди презентації

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Атестаційна робота магістра



Дослідження методів фільтрації даних в соціальних мережах

Керівник, проф.

І.В. Шостак

Виконала

ст. гр. ІПЗмзд-18-1

Я.І. Забіяка

1

Мета роботи



Аналіз методів агрегації даних для створення необхідних пакетів в даних користувача соціальних мереж

- Агрегація даних на основі веб-платформи включає агрегування загальнодоступних даних про людину з веб-сайтів соціальних мереж, таких як Facebook, LinkedIn і т.д.
- Платформа може взаємодіяти з постачальниками соціальних мереж, щоб отримати дані в режимі реального часу на основі імені особи і змісту профілю.

2



Актуальність

- Сучасні соціальні мережі не допомагають користувачам організувати пакети інформації. Вони вимагають різних облікових даних для входу в різні соціальні мережі, це не є зручним оскільки користувач має запам'ятати кілька логінів, паролів, а також запам'ятовувати перелік друзів які використовують ту чи іншу соціальну мережу.
- Агрегатори соціальних мереж –відносно новий вид додатків, за допомогою яких намагаються консолідувати всі свої різні профілі соціальних мереж в одне ціле.

3



Інструменти персоналізації пошуку в ІПС

Існує три види пошукових систем, які можуть забезпечити персоналізовану інформацію:

- системи зворотного зв'язку по релевантності запиту;
- системи демографічної інформації;
- системи, що базуються на аналізі попередніх запитів користувачів.

Під персоналізацією пошуку розуміється надання користувачеві індивідуальних, персоналізованих результатів залежно від його інформаційних потреб, пріоритетів, інтересів, географічного положення, соціального стану, віку та інших особливостей.

4

Методи, що використовуються для персоналізації пошуку.



- Методи, засновані на неявному зворотному зв'язку з користувачами - дані методи орієнтовані на аналіз контекстної інформації та поведінки користувача у фоновому режимі при роботі з ІПС
- Методи, засновані на явному зворотному зв'язку з користувачами - у цьому випадку персоналізація пошуку виконується на основі аналізу даних профілю користувача, зазначених при реєстрації, переданих користувачем даних в ІПС і зібраної статистики
- Існує зростаюча тенденція до використання інструментів моніторингу соціальних медіа, які дозволяють маркетологам і компаніям шукати, відстежувати і аналізувати бесіди он-лайн про їх торгову марку, продукти, про теми, що пов'язані з їхнім бізнесом і що для них представляє інтерес.

5

Елементи «Сотової структури»



- **ідентичність** – ступінь, в якій користувачі розкривають свою особистість в умовах соціальних медіа.
- **розмови** – ступінь спілкування користувача з іншими користувачами в соціальних медіа.
- **спільне використання** – представляє собою ступінь, в якій користувачі обмінюються, поширюють і отримують контент,
- **наявність** – представляє собою ступінь доступності користувача як в реальному так і віртуальному просторі.
- **відносини** – представляє собою ступінь зв'язку користувачів;
- **репутація** – представляє собою ступінь, в якій користувачі можуть ідентифікувати положення інших людей, в тому числі себе, в соціальних медіа.
- **групи** – представляє собою ступінь, в якій можуть утворювати спільноти користувачів і суб-спільноти людей за будь-якими критеріями

6

Огляд і аналіз соціальних медіа



- соціальні медіа даних, за типами даних;
- соціальні медіа програмні послуги доступу до даних та інструменти для пошуку (текстових) даних із соціальних мереж засобів масової інформації,
- RSS-канали, новини і т.д.
- джерела даних, послуги і інструменти ;
- дані каналів за допомогою API-інтерфейсів;
- інструменти аналізу тексту;
- прості інструменти;
- інструменти аналізу;
- соціальні медіа-платформи
- новинні платформи.

7

Постановка задач дослідження



Он-лайн механізм пакетів даних в подальшому використовуватиме аналітик, маркетолог, менеджер для звітності, обробки інформації тощо.

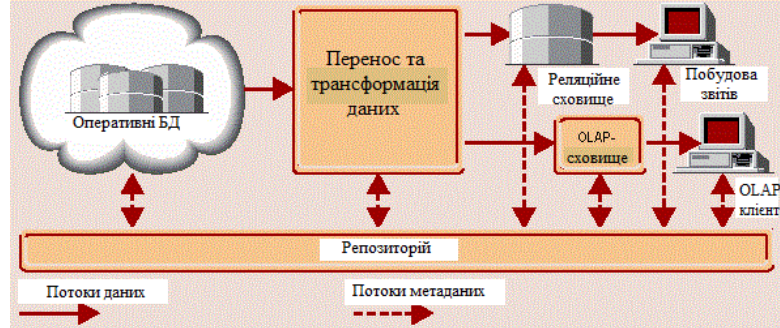
Це дозволить аналітикам і менеджерам електронного бізнесу отримувати уявлення про користувача завдяки отриманій інформації та за допомогою

Для досягнення зазначеної мети необхідно розв'язати наступні завдання:

- вивчення основних характеристик і особливостей ІПС, а також різних методів пошуку в ІПС;
- аналіз проблем побудови ІПС Інтернету та можливих шляхів їх рішення;
- розробка методів інтелектуалізації та персоналізації ІПС;
- розробка архітектури інтелектуальної ІПС на основі прецедентів;
- програмна реалізація прототипу ІПС на основі прецедентів

8

Аналіз структур сховищ даних

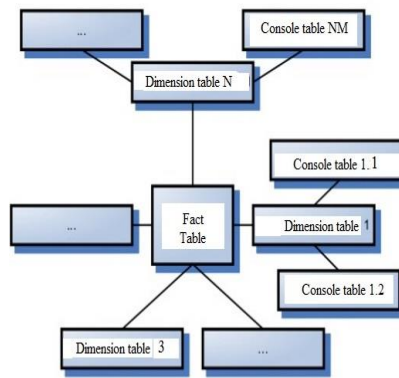


9

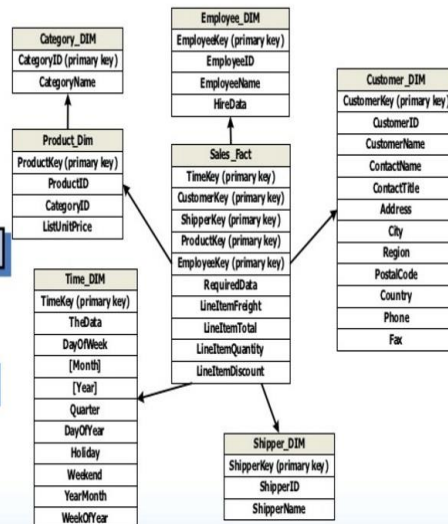
Логічне та фізичне представлення схеми «Сніжинка»



Логічне представлення схеми "Сніжинка"

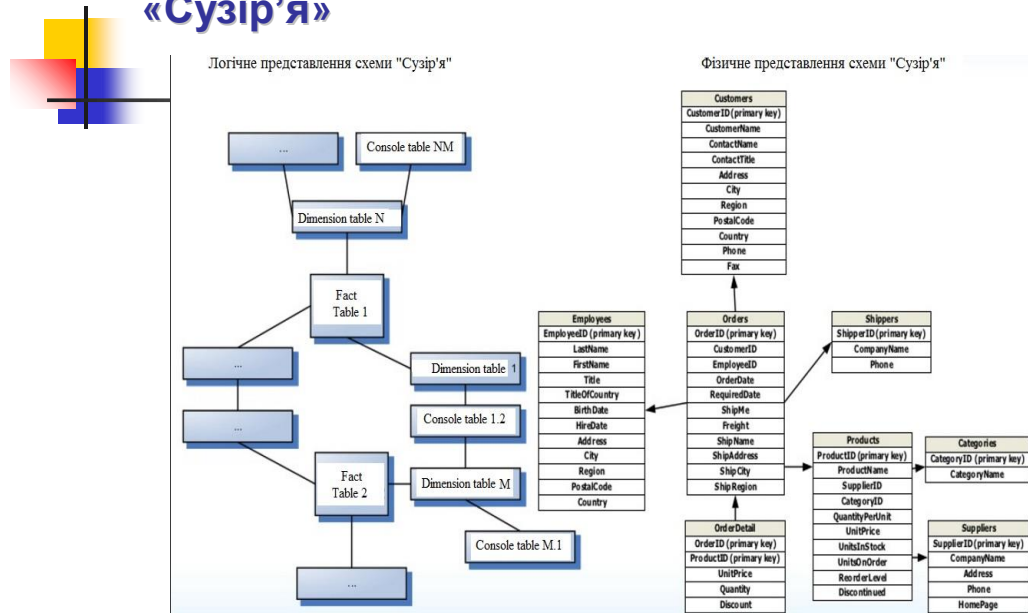


Фізичне представлення схеми "Сніжинка"



10

Логічне та фізичне представлення схеми «Сузір'я»



11

Стек технологій для аналізу даних користувача з соціальних мереж містить

- фреймворк для збору реальних призначених для користувача даних шляхом звернення до веб-інтерфейсів соціальних сервісів;
- інструмент для генерації випадкових соціальних графів із заданими структурними властивостями;
- методи обробки текстових даних користувачів соціальних мереж: визначення демографічних атрибутів шляхом лінгвістичного аналізу профілів і текстів повідомлень, а також пошук описів подій в повідомленнях;
- методи обробки мережевих даних (соціальних зв'язків між користувачами), а саме метод ідентифікації користувачів різних соціальних мереж та метод пошуку співтовариств користувачів;
- метод вимірювання інформаційного впливу і пошуку найбільш впливових користувачів.

12

Фреймворк для збору даних з різних інтернет-сервісів



Реалізовано кілька способів отримання репрезентативних вибірок користувачів соціальних мереж:

- семплірованіє методом обходу в ширину (пошук в ширину, breadth-first search BFS),
- по Метрополісу-Гастінгс (Metropolis-Hastings Random Walk – MHRW)
- методом «лісової пожежі» (Forest Fire -FF).

Реалізація механізму автоматичного вибору облікового запису соціальної мережі для кожного запиту, а також підтримка проксі-з'єднань, що забезпечує стійкість до блокувань по IP-адресам облікових записів. Крім того, фреймворк підтримує багатопоточне скачування.

Однією з ключових особливостей такого фреймворку є можливість швидко реалізувати нові сценарії скачування і методи семплінгу.

13

Фреймворк для збору даних



Математична модель агрегації для узагальнення графа

Пропонується модель узагальнення графа (МУГ) на основі угруповання вузлів на атрибутах і парних відносин, це дає короткий граф вхідного графа за допомогою угруповання вузлів на основі вибраних користувачем атрибутів вузлів і зв'язків.

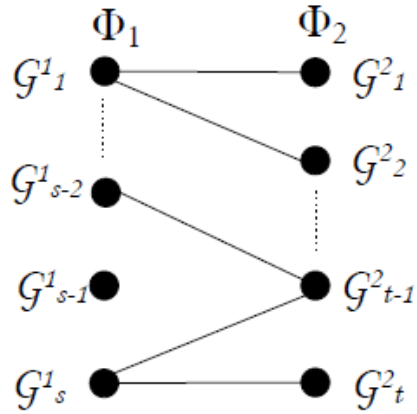
- Різні типи ребер в графі відповідають різним типам відносин між вузлами
- Операція узагальнення графа на основі угруповання вузлів на атрибутах і парних відносин виробляє короткий граф через однорідне угруповання вузлів вхідного графа, засноване на обраних користувачем атрибутах вузлів і зв'язках

14

Побудова двочасткового графу $\Phi_1 \cup \Phi_2$.



Двочастковий граф Φ_3



15

Операція УГОУВ



приймає в якості вхідних даних граф G , набір атрибутів $A \subseteq \Lambda(G)$

і набір ребер $R \subseteq \Upsilon(G)$

і отримується сумарний узагальнюючий граф G_{total}

- Таким чином, операція УГОУВ виробляє узагальнений граф вхідного графа на основі вибраних користувачем атрибутів і зв'язків.
- Вузли цього короткого графа відповідають групам в максимумі (A, R) , що сумісні з групуваннями.
- Ребра цього короткого графа є групові відносини виведені з вузла відносин в R .
- Отже представлена операція агрегації УГОУВ заснована на угрупованні графа.

Цей метод дозволяє користувачам вільно вибирати атрибути вузлів і відносини, які становлять інтерес, і виробляють угруповання на основі певних функцій.

16

Запит для доступу до властивостей об'єкта

Для того, щоб отримати, його унікальний ідентифікатор повинен бути відомий, щоб зробити виклик API. Пошук API Facebook (частина Graph API Facebook,) можна отримати за запитом:

<https://graph.facebook.com/search?q=QUERY&type=page> .
Докладний формат API запиту показаний нижче.

```
// 2. Приклад Facebook Graph API Search Query Format
GET graph.facebook.com
/search?
q={your-query}&
[type={object-type}]{#searchtypes}
```

Тут, «QUERY» може бути замінений будь-яким терміном пошуку, а «page» може бути замінена на «post», «user», «page», «event», «group», «place», «checkin», «location» or «placetopic».

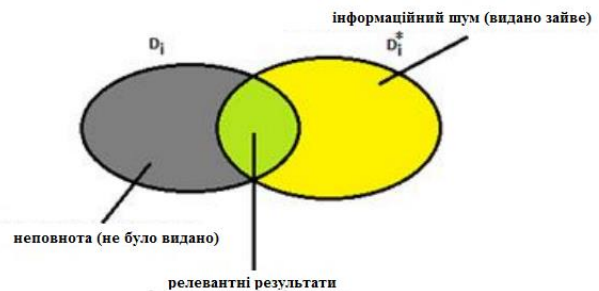
Результати цього пошуку будуть містити унікальний ідентифікатор для кожного об'єкта.

17

Оцінка інтегрального показника ефективності інформаційного пошуку

k_{Π} – коефіцієнт інформаційної повноти;
 $k_{\text{ш}}$ – коефіцієнт інформаційного шуму.

$$k_{\Pi} = \frac{|D_i \cap D_i^*|}{|D_i|}, \quad k_{\text{ш}} = \frac{|D_i^* \setminus D_i|}{|D_i^*|}.$$



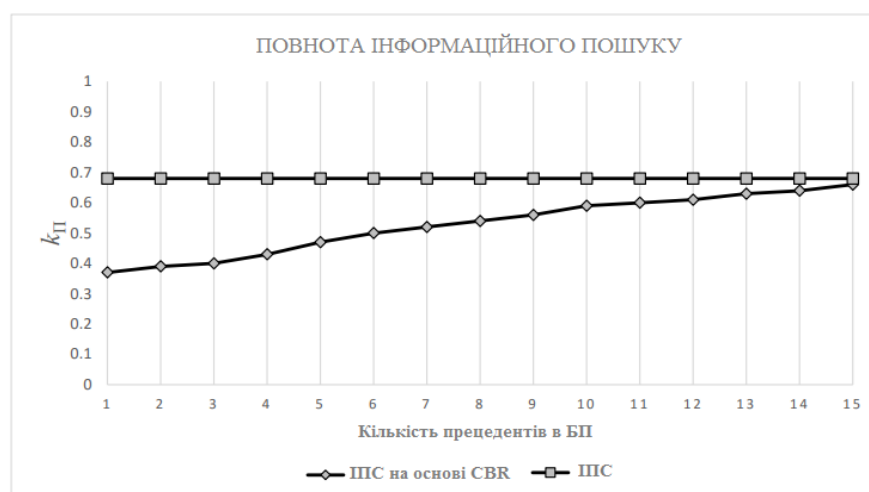
18

Значення показників ефективності для стандартної ІПС і ІПС на основі розроблених технологій

	ІПС	1 прецедент у БП CBR-агента	2 прецедент у БП CBR-агента	3 прецедент у БП CBR-агента	4 прецедент у БП CBR-агента	5 прецедент у БП CBR-агента	6 прецедент у БП CBR-агента	7 прецедент у БП CBR-агента	8 прецедент у БП CBR-агента	9 прецедент у БП CBR-агента	10 прецедент у БП CBR-агента	11 прецедент у БП CBR-агента	12 прецедент у БП CBR-агента	13 прецедент у БП CBR-агента	14 прецедент у БП CBR-агента	15 прецедент у БП CBR-агента
K_{Π}	0,68	0,37	0,39	0,40	0,43	0,47	0,50	0,52	0,54	0,56	0,59	0,60	0,61	0,63	0,64	0,66
$K_{\#}$	0,33	0,17	0,17	0,18	0,19	0,21	0,22	0,23	0,23	0,24	0,25	0,25	0,25	0,26	0,27	0,29
E_1	0,70	0,51	0,52	0,53	0,55	0,57	0,69	0,60	0,60	0,61	0,63	0,63	0,64	0,65	0,65	0,67

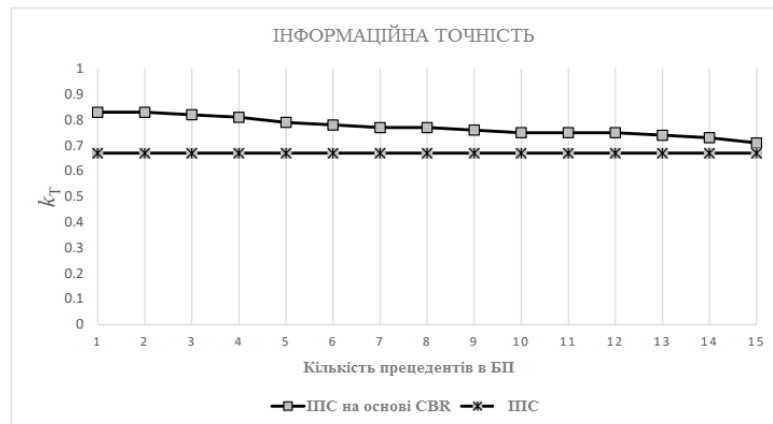
19

Значення k_{Π} для ІПС з використанням CBR-агентів і без використання CBR-агентів



20

Значення k_T для ІПС з використанням CBR-агентів і без використання CBR-агентів



Використання запропонованої архітектури для інформаційно-пошукової системи припускає, що результат відповіді на запит може бути отриманий у результаті взаємодії агентів

21

Висновки



В роботі було проведений аналіз методів агрегації даних для створення необхідних пакетів даних користувача соціальних мереж з використанням підходу Інтернет аналітичної обробки даних, як простого типу агрегації даних для створення необхідних пакетів даних користувача соціальних мереж.

Наведено аналіз методів для організації сховища даних, платформ агрегації даних у соціальних мережах.

Детально представлена практична реалізація методів агрегації даних в соціальних медіа. Всі програмні приклади наведено для JSON

В роботі запропоновані:

моделі пошуку співтовариств користувачів та інтеграції соціальних даних через сайти соціальних мереж. Ці моделі використовуватимуться для вирішення проблем відстежування соціальної активності користувачів соціальних мереж та перевантаження великою кількістю соціальних даних;

математична модель агрегації для узагальнення графа. Математично визначена операція агрегації узагальнення графа на основі угруповування вузлів на атрибутах і парних відносин (УГОУВ) виробляє короткий граф через однорідне угруповування вузлів вхідного графа, заснований на обраних користувачем атрибутах вузлів і зв'язках, що засновано на угрупованні графа.

Цей метод дозволяє користувачам вільно вибирати атрибути вузлів і відносини, які становлять інтерес, і виробляють угруповання на основі певних функцій.

22

ДОДАТК В

Апробація результатів дослідження

Підготовлено тези доповіді на Міжнародний молодіжний форум

«Радіоелектроніка і молодь в ХХІ сторіччі»

ДОСЛІДЖЕННЯ МЕТОДІВ ФІЛЬТРАЦІЇ ДАНИХ В СОЦІАЛЬНИХ МЕРЕЖАХ

Зайцева Н.І.
Научний співробітник – проф. Шостак І.В.
Львівський національний університет радіоелектроніки
(81166, Львів, пр. Науки, 14, к.а.ф. Програмної інженерії,
тел.: (037) 702-4-46)
E-mail: i.v.shostak@gmail.com

In the work the method of estimating the maturity (perfection) of the projects performed in the testing of software systems is developed. This method is based on TDm's five-level maturity model. Also, the use of mathematical model, the basic process of testing software data processing systems under the conditions of limited resources, as well as several methods aimed at improving the quality of software products, have been developed and substantiated.

До головних напрямків програмної інженерії відносяться задачі вдосконалення процесів життєвого циклу ПС, зокрема процесу тестування. Неупереджене оцінювання якості програмних продуктів – основна мета програмної інженерії та прамеє судило розробки ПС.

Відповідь на питання, як підвищити конкурентоспроможність українських програмних продуктів, як знизити ризик проєкту, як досягти балансу сторін трикутника «якість – вартість – час проєкту», програмом остаточно розв'язується завдяки відповідь не лише керівники організацій-розробники ПС і менеджери проєктів, але і замовники і споживачі, що використовують програмні продукти низької якості. Наслідки постачання програмних продуктів низької якості – це завжди збитки користувачів не лише матеріальні і фінансові втрати, але і психічні страждання. Ці проблеми в свою чергу, негативно відображаються на конкурентоспроможності організацій-розробників програмних продуктів. Для забезпечення необхідного рівня якості ПС в міжнародній практиці знаходять застосування два підходи: продукто-орієнтований і процесно-орієнтований. В першому випадку робиться на оцінку якості шляхом тестування готового програмного продукту. Цей підхід базується на припущенні, що чим більше виявлено і усунуто дефектів в ПС при тестуванні, тим вище його якість.

Тестування – важливий етап розроблення ПС, оскільки, з одного боку, вимагає значних витрат на проведення, а з іншого – робить великий внесок у його якість. Через гуртовий доведену неможливість вичерпного тестування та велику поточну вартість витрат через відмови в ПС, потрібен чітко визначений та ефективний процес тестування, базований на ухваленні об'єктивних рішень щодо тривалості та вартості тестування для досягнення необхідного рівня довіри до якості ПС.

Метою визначення кількісних критеріїв завершення тестування та керування процесом тестування і використання кількісних вимірів ще мало використовуються в проєкті створення ПС, що призводить до того, що якість та надійність захищеної не передбачуваними. Лише за наявності достовірної та своєчасної інформації щодо стану ПС, видів ризиків і можливих витрат через відмови може бути забезпечене ефективне виконання процесу тестування.

Метою роботи є проведення комплексу досліджень з інженерії тестування ПС: оброблення даних, формування ефективної стратегії тестування програмних систем, спрямованої на визначення рівня відмов під час експлуатації. Для цього в роботі розв'язуються наступні завдання:

дослідження сучасних вимог до процесу тестування ПС;
аналіз існуючих моделей надійності ПС, розроблення алгоритмів та програм для реалізації;

побудова моделей визначення оптимального часу тестування модулів ПС і урахування ризиків відмов;

визначення структури базового процесу, що регламентує всі дії з підготовки, проведення та оцінювання результатів тестування, та розроблення методик виконання процесу тестування ПС оброблення даних;

проєктування та реалізація програмного комплексу підтримки інженерії тестування;

аналіз сучасних підходів до визначення рівня якості процесу тестування ПС та розроблення методик оцінювання процесу тестування, запровадження запропонованих підходів, моделей та методик інженерії тестування в проєкті і розроблення ПС оброблення даних та опис результатів випробування запропонованих моделей оцінювання оптимального часу тестування та методу оцінювання ризиків відмов програмних модулів.

З метою визначення ефективності запровадження базового процесу тестування був розроблений метод оцінювання якості (доповідності) виконуваних у проєкті дій з тестування ПС. Цей метод ґрунтується на підтриманні моделі якості TDm.

Випробування моделей мають використовувати в комерційній інформаційно-аналітичній системі підтримки прийняття управлінських рішень, яка має складатися з програмних компонентів, об'єднаних єдиним процесом оброблення даних. В кожній практичній реалізації аналіз витрат відмов системи показав, що з поточної кількості інформації найбільший внесок у ризик її відмов робить ПК контролю і введення даних до БД ORACLE, який використовується та основною функцією на її робочих місцях. Для того, щоб модуль цього ПК був виконаний оцінкою ризиків відмов та часу тестування.