

ПРАКТИЧНА РОБОТА № 4

Тема роботи: Захист веб-ресурсів від ботів та спаму за допомогою механізму CAPTCHA.

Мета роботи: отримати практичні навички реалізації механізму CAPTCHA.

Теоретичні відомості CAPTCHA (англ. «completely automated public turing test to tell computers and humans apart» повністю автоматизований публічний тест Тюринга для розрізнення комп'ютерів і людей) - торгова марка Університет Карнегі - Меллона, комп'ютерний тест типу виклик-відповідь, який використовується для того, щоб визначити, хто використовує систему - людина чи комп'ютер

У найпоширенішому варіанті CAPTCHA від користувача потрібно ввести символи, зображені, як правило, в спотвореному вигляді на пропонованому малюнку, іноді з додаванням шуму або напівпрозорості.

Рідше застосовуються CAPTCHA, що базується на розпізнаванні мови (в основному як альтернатива для людей з порушеннями зору), або на інших варіантах завдань штучного інтелекту.

CAPTCHA найчастіше використовується при необхідності запобігти використанню інтернет-сервісів ботами, зокрема, для запобігання автоматичній реєстрації, викачуванню файлів, масовим розсилкам тощо.

Для чого потрібний тест CAPTCHA? У сучасному світі людині все частіше доводиться зіштовхуватися з роботами. Роботи керують трубопроводами, підтримують технологічний процес, виконують складні обчислення. Існує безліч роботів і в мережі Internet - тут їх, як правило, скорочено називають ботами. Internet боти стежать за порядком у чатах, займаються розсиланням листів, індексую WEB сторінки для прискорення пошуку. Однак не всі боти займаються корисною діяльністю. Існує безліч ботів написаних для деструктивних цілей.

На просторах мережі часто зустрічаються:

- боти для захватів чат - кімнат в IRC;
- боти для збору адресів e-mail і розсилання за ними спаму (небажаної кореспонденції, як правило рекламного характеру);
- флуд-боти, що займаються —засміченням форумів і гостьових книг;
- боти, що проводять масову реєстрацію облікових записів на поштових й інших сервісах;
- боти, що займаються автоматичним підбором паролів;
- боти, що роблять накручення лічильників відвідуваності сайтів;

- боти для розсилання SMS через WEB інтерфейси;
- боти для участі в голосуваннях;
- й інші.

Рішенням проблеми ботів може стати тест CAPTCHA. Цей тест найчастіше застосовується для захисту WEB форм від автоматичного заповнення. Іноді дуже важливо чи знати заповнена форма людиною або ж її заповнив бот.

Людину від бота, звичайно, можна відрізнити за рядом непрямих ознак наприклад по швидкості заповнення форми (комп'ютер робить це набагато швидше людини) або по IP адресу (з одного IP адресу надходить цілий ряд запитів на заповнення однієї й тієї ж форми) та ін.

Такі перевірки робляться досить простими методами. Але за простотою таких перевірок криється й простота їхнього обходу з боку бота. Тому такі непрямі перевірки застосовуються лише як допоміжний засіб, а як основний засіб виступає саме CAPTCHA.

Варіанти реалізації CAPTCHA.

На сьогоднішній день існує безліч реалізацій цього тесту для мережі Internet від примітивних до найскладніших, які тільки здатна видумати людина.

Перелічимо деякі з них:

1. Найпоширеніша реалізація CAPTCHA - це буквено-цифровий тест.

Суть його полягає у наступному: при заповненні форми людині показується картинка з деяким випадковим набором букв та цифр, а поруч із картинкою текстове поле в яке необхідно ввести символи зображені на картинці. Якщо введені в текстовому полі символи й символи на картинці збігаються тоді заповнена форма приймається до обробки, а якщо ні, то генерується нова картинка й людини просять увести ще раз символи з картинки. Це просте завдання для людини, але для бота воно вже досить важке. Як правило зображення на картинці зашумлюють для ускладнення спроб автоматичного розпізнавання.

Приклад цієї реалізації можна побачити на рисунку 1.



Рисунок 1 - Варіант реалізації CAPTCHA № 1

2. Іноді застосовують спрощену для користувача реалізацію описаного в першому пункті тесту, коли користувачеві показують усе той же випадковий набір символів і пропонують кілька варіантів відповіді. У цьому варіанті реалізації тесту вже неможливо зробити помилку при наборі тексту й тому після першої або другої неправильної відповіді IP адреса блокується на кілька годин. Приклад цієї реалізації можна побачити на рисунку 2.

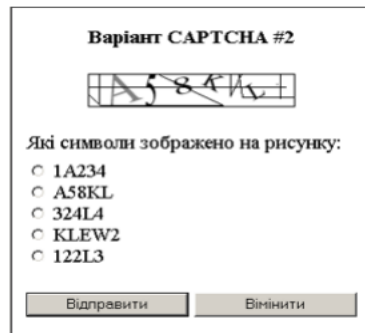


Рисунок 2 - Варіант реалізації CAPTCHA № 2

3. Існує також комбінація першого й другого варіанта тесту. Людині виводиться у вигляді картинки текст, що спонукує його до деякої дії (наприклад —натисніть на хвіст кішці). А поруч із текстом розташовується картинка з декількома об'єктами (у нашому прикладі ними можуть бути кішка, собака, равлик). Координати натискання по картинці відслідковуються й перевіряються. Причому питання й супутня йому картинка із предметами щоразу різні. Приклад цієї реалізації можна побачити на рисунку 3.



Рисунок 3 - Варіант реалізації CAPTCHA № 3

4. Людині показують картинку з якимось об'єктом (наприклад портрет Шевченка або Франка) і просять увести його назву (у нашому випадку прізвище письменника). Як правило картинки перевернуті випадковими 24 дефектами, щоб бота не навчили розпізнавати всі картинки тесту. Приклад цієї реалізації можна побачити на рисунку 4.

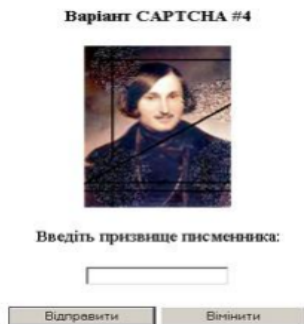


Рис. 4. Варіант реалізації CAPTCHA № 4

5. Людині показують кілька картинок (частіше всього 3 або 4), на яких зображено декілька предметів, і просять ввести назву того предмета, що є на всіх картинках.

Приклад цієї реалізації можна побачити на рисунку 5.

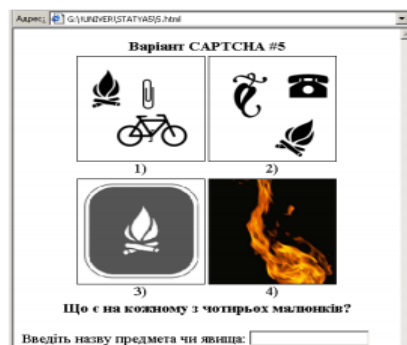


Рисунок 5 - Варіант реалізації CAPTCHA № 5

6. Математична CAPTCHA вона є різновидом буквено-цифрової, але замість букв і цифр зображених на малюнку потрібно підрахувати та ввести результат математичної дії.

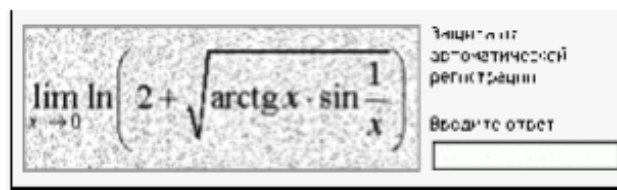


Рисунок 6 - Варіант реалізації CAPTCHA № 6

Є й інші варіанти реалізації CAPTCHA, проте вони досить складні у програмній реалізації і тому використовуються дуже рідко.

Завдання для виконання

Удосконалити ПЗ розроблене в попередніх практичних роботах в частині захисту однієї із розроблених форм за допомогою механізму CAPTCHA відповідно до індивідуального завдання.

Індивідуальне завдання

1. Для оцінки «задовільно» запрограмувати механізм CAPTCHA у вигляді запиту відповіді на звичайне зображення, яке в довільному порядку обирається з бази даних ПЗ.
2. Для оцінки «добре» запрограмувати механізм CAPTCHA у вигляді запиту відповіді на звичайне зображення із накладенням шумів, яке в довільному порядку обирається з бази даних ПЗ. Розмістити на окремій формі reCaptcha Google.
3. Для оцінки «відмінно» запрограмувати механізм CAPTCHA у вигляді запиту відповіді на зображення із накладенням фону та шумів, яке в довільному порядку обирається з бази даних ПЗ. Розмістити на окремій формі reCaptcha Google.

Зміст звіту

4. Тема та мета роботи.
5. Завдання на виконання.
6. Опис функцій розробленого програмного забезпечення.
7. Програмний код, що реалізовує індивідуальне завдання.
8. Тестовий приклад, що відображає результат виконання практичної роботи.
9. Висновок

Контрольні запитання

1. Історія створення CAPTCHA.
2. Механізми створення CAPTCHA.
3. Механізми «обходу» CAPTCHA
4. Як CAPTCHA допомагають боротись із спамом.
5. Переваги використання CAPTCHA.
6. Варіанти реалізації CAPTCHA.
7. Особливості використання reCaptcha Google.
8. Особливості передбачення результатів CAPTCHA.
9. Переваги використання CAPTCHA.
10. Відомі CAPTCHA-сервіси.