Politechnika Wrocławska

Wydział Informatyki i Telekomunikacji

Obliczenia Wysokiej Wydajności

Środa, 13:15-15:00

Łamanie zahaszowanych haseł metodą słownikową

Autorzy:

Aleksandra Rozmus, 252954

Paweł Struczewski, 252950

Prowadzący:

dr inż. Radosław Idzikowski

18 października 2023

Politechnika
Wrocławska

Spis treści

1 Temat projektu

Tematem projektu jest stworzenie wielowątkowego współbieżnego programu do łamania zahaszowanych haseł metodą słownikową z wykorzystaniem algorytmu MD5.

2 Skład zespołu

Zespół składa się z dwóch osób:

- Aleksandra Rozmus - 252954
- Paweł Struczewski - 252950

3 Zakres projektu

Podczas tego projektu stworzony zostanie program polegający na łamaniu zahaszowanych haseł. Hasła są generowane metodą skrótów kryptograficznych MD5. Program zawiera listy zaszyfrowanych haseł do łamania oraz słownik, który zawierać będzie dużą liczbę słów różnej długości. Lista haseł będzie typowo mniejsza (maksymalnie 1000 jednowyrazowych haseł) i będą one miały stałą długość 32 znaków.

Zaimplementowani producenci haseł będą tworzyli hasła w następującym formacie:

- wyłącznie małe litery
- pierwsza wielka litera i reszta małe litery
- wyłącznie wielkie litery

Do każdego hasła przed zahaszowaniem dodana może zostać cyfra na początku wyrazu, na końcu lub w obu wymienionych miejscach.

Każde kolejno znalezione złamane hasło powinno zostać wyświetlone na ekranie.

Projekt podzielony został na 3 etapy:

- Etap 1 - Implementacja metody sekwencyjnej
- Etap 2 - Implementacja metody równoległej
- Etap 3 - Badania - porównanie metody sekwencyjnej i równoległej

W ramach implementacji metody równoległej stworzony zostanie wątek główny, którego zadaniem jest wczytanie do pamięci globalnej programu podanej przez użytkownika listy zaszyfrowanych haseł do łamania, oraz podanego (lub domyślnego) słownika. Następnie wątek główny powinien uruchomić zestaw wątków "producentów". Każdy z tych producentów ma za zadanie próbować złamać wszystkie hasła, stosując różne metody budowy haseł, aby proces łamania haseł mógł działać równolegle. Każdy producent będzie kolejno tworzyć potencjalne hasła przy użyciu swojej unikalnej metody, a następnie obliczać 32-znakowy skrót kryptograficzny MD5 dla każdego z tych haseł. Następnie porówna ten skrót z całym zestawem zaszyfrowanych haseł. Jeśli którykolwiek z tych zaszyfrowanych haseł zgadza się z obliczonym skrótem, to oznacza, że odnaleziono oryginalne hasło. To hasło zostanie przekazane do wątku konsumenta, który je zarejestruje i oznaczy odpowiednie zaszyfrowane hasło w globalnej tablicy, aby uniknąć dalszych prób jego porównywania. Wątek konsumenta powinien wyświetlać na ekranie każde kolejno znalezione złamane hasło w miarę otrzymywania ich od producentów.

Ilość działających wątków będzie dostosowywana do ilości rdzeni maszyny, na której uruchomiony jest program.

4 Wyliczenia teoretycznego przyśpieszenia

Teoretyczne przyśpieszenie programu wielowątkowego w zadaniu łamania zahaszowanych haseł za pomocą metody słownikowej może być wyliczone przy użyciu tzw. prawa Amdahla.

Wzór wykorzystany do wykonania obliczeń:

$$Pr = \frac{1}{(1-P) + \frac{P}{N}}$$

gdzie:

P - Proporcja programu, która może podlegać zrównolegleniu,
N - liczba procesorów.

W naszym przypadku. Cześć nie podlegająca zrównolegleniu to około 20%. Jest to część odpowiedzialna za np. operacje wczytywania i porównywania. Więc część programu podlegająca zrównolegleniu to około 80%. Przykładowe obliczenia dla komputera z dostępnymi 4 rdzeniami:

$$Pr = \frac{1}{(1-0,8) + \frac{0,8}{4}} = 2,5$$

Przyśpieszenie równoległe wyniesie około 2,5 dla komputera z dostępnymi 4 rdzeniami.