Admis la susținere
Șef departament: Fiodorov I. dr., conf.univ.

„___" _____ 2022

Analiza traficului prin algortimi de învățare automată și metode de îmbunătățire a lui

Teză de master în

Student: _____ **Gaidău Mihai, TI-201M**

Coordonator: _____ **Catruc Mariana, lect. univ.**

Consultant: _____ **Cojocaru Svetlana, lect. univ.**

Chișinău, 2022

Rezumat

Soluția tehnică din această teză de master, de analiză a traficului prin algortimi de învățare automată și metode de îmbunătățire a lui este destinată pentru a veni în ajutorul statului Republica Moldova și alte state care au nevoie de un aplicare a unei soluții digitale bazate pe un algoritm care are ca scop analiza unor parametri critici. Astfel, parametrul de bază al algormitmilor care sunt vizați în această lucrare este fluxul de autovehicule și ca rezultat este așteptat un timp cât mai minim de parcurge a tuturor orașelor.

Acest proiect începe prin etapa de identificare a problemei și continuă cu etapa secundară de analiză a domeniului de aplicare a algoritmilor definiți. Astfel în timpul cercetării soluția tehnică se v-a analiza, compara cu soluții deja existente, astfel aceasta v-a pune în evidență aspectele puternice ale acesteia dar și acelea asupra cărora este necesar de îmbunătățit. Mai mult ca atât algoritmii de optimizare pot fi implementatți nu doar în problema descrisă dar și alte tipuri de probleme în care trebuie de aflat cel mai scurt drum dintre n orașe, cel mai optimizat traseu și cel mai puțin timp pentru a parcurge o anumită distanță.

Cu scopul de a defini cât mai explicit ceea ce presupune această soluție, rezultatele în baza parametrilor descriși de intrare sunt afișate în diagrame, grafice și tabele care sunt pentru compararea și identificarea celui mai optim algoritm. În cadrul acestui proiect modulul este unul singular și cel mai important, deoarece aplicația este monolitică în care fiecare algoritm este stocat într-un fișier. Datorită naturei problemei care este abordată și soluției propuse, se înțelege că este nevoie de un limbaj de programare cu putere de calcul destul de mare și viteza de procesare la fel, ridicată. Astfel, C++ este un limbaj de nivel jos care operează cu datele și calculele rapid, în timp optim pentru a fi posibilă analiza și implementarea acestora.

Nu în ultimul rând trebuie de luat în calcul complexitatea acestor algoritmi, deoarece, acest criteriu este unul de bază și cel care explică de ce un algoritm a executat mai repede un calcul și altul mai mult timp. În cazul algoritmilor implementați se observă complexitate algoritmică, factorială și exponențială. Prin intermediul graficelor acestea este ușor vizibil.

Etapele de realizare și implementare a soluției tehnice vin să asigure identificarea celui mai optim traseu dintre n orașe respectând condiția de a parcurge pe toate menționate. Odată cu etapa finală vine și analiza complexă a algoritmului care afișează cel mai bun rezultat, cel mai puțin timp de parcurgere și aplicarea acestuia în problema definită de către societate.

Abstract

The technical solution in this master's thesis, traffic analysis through machine learning algorithms and methods to improve it is intended to come to the aid of the Republic of Moldova and other states that need an application of a digital solution based on an algorithm. which aims to analyze some critical parameters. Thus, the basic parameter of the algorithms that are targeted in this paper is the flow of vehicles and as a result, a minimum time to travel through all cities is expected.

This project starts with the problem identification stage and continues with the secondary analysis stage of the scope of the defined algorithms. Thus, during the research, the technical solution was analyzed, compared with existing solutions, so it highlighted the strong aspects of it but also those that need to be improved. Moreover, the optimization algorithms can be implemented not only in the described problem but also in other types of problems in which the shortest route between n cities, the most optimized route, and the least time to cover a certain distance must be found.

To define as explicitly as possible what this solution entails, the results based on the described input parameters are displayed in diagrams, graphs, and tables that are for comparing and identifying the best algorithm. In this project the module is unique and the most important, because the application is monolithic in which each algorithm is stored in a file. Due to the nature of the problem being addressed and the proposed solution, it is understood that a programming language with fairly high computing power and a similarly high processing speed is needed. Thus, C ++ is a low-level language that operates with data and calculations quickly, in the optimal time to be able to analyze and implement them.

Last but not least, the complexity of these algorithms must be taken into account, because this criterion is a basic one and the one that explains why one algorithm performed one calculation faster and another longer. In the case of the implemented algorithms, algorithmic, factorial, and exponential complexity are observed. It is easily visible through the graphics.

The stages of realization and implementation of the technical solution come to ensure the identification of the best route between n cities respecting the condition to cover all mentioned. Along with the final stage comes the complex analysis of the algorithm that displays the best result, the least time to travel, and its application in the problem defined by the company.

Table of Contents

INTRODUCTION

Life evolves, and it gets more and more technologically advanced which means to make life easier. From the first computer to the current supercomputers, man made huge progress and made this power to be accessible from everywhere. Nowadays computers are all around, doing a lot of types of jobs, from making calculations that measure the size of the universe, to providing a platform where to write this thesis.

By using computer power man is trying to resolve all types of problems from different domains. This report will approach a major problem that appears to be frequent in fast-growing cities and cities with a bad infrastructure – the traffic problem. Traffic became a big problem since it's directly related to the economy.

The solution to this difficult problem needs to have a good architecture, to be easily scalable and it has to use a lot of technologies to make a good result. For having a good application, with impact over-described problem, is needed to analyze the statistical data and history. In terms of traveling problems or how to access all the cities from a country in the most optimal time, choosing the best road, and keeping the path to the initial city or point, daily monitoring and algorithms can help.

As implementing such a major system needs a lot of work, it was decided to simplify the problem by relating it to a close mathematical model – the travel-salesman problem. So, starting with the Hamiltonian cycle that identifies a tour that visits every city exactly once. The main scope of this cycle is to find the minimum weight. Implementation of Hamiltonian needs to apply dynamic programming and take into account hardware characteristics because it performs calculations and operates with memory.

Next, the second algorithm, TSP defines the challenge of finding the shortest yet most efficient route for a person to take given a list of specific destinations. It is a known algorithmic problem in the fields of computer science and operations research. TSP has commanded so much attention because it's so easy to describe yet so difficult to solve. TSP belongs to the class of combinatorial optimization problems known as NP-complete. This means that TSP is classified as NP-hard because it has no "quick" solution and the complexity of calculating the best route will increase when you add more destinations to the problem.

The travel salesman problem can be solved by analyzing every round-trip route to identify the shortest one. So, logically if the number of destinations increases the capabilities should be scaled and the hardware should be more advanced. There are a lot of commercial solutions that can solve and implement algorithms in each case, but for this implementation need to pay a huge amount of money. For better performing and customization of the application of the best algorithm, it is needed to implement my own.

BIBLIOGRAPHY

1. Muneera C P, Krishanamurthy Karuppanagouner, *Economic Impact of Traffic Congestion – Estimation and Challenges* [online]. June, 2008 [cited: 10.09.2021]
   Available: (PDF) Economic impact of traffic congestion- estimation and challenges (researchgate.net)

2. Francesca Backer, *The technology that could end traffic* jams [online] December 2018, [cited: 15.09.2021]
   Available: https://www.bbc.com/future/article/20181212-can-artificial-intelligence-end-traffic-jams

3. Corinne Brucato, THE TRAVELING SALESMAN PROBLEM, B.A., Sonoma State University, 2010, M.S., University of Pittsburgh, 2013 [online], 2013, [cited: 02.10.21]
   Available: The Traveling Salesman Problem (pitt.edu)

4. Borwein, J. M. and Borwein, *P. B. Pi & the AGM: A Study in Analytic Number Theory and Computational Complexity.* New York: Wiley, Published: 1987, [cited 15.10.2021].

5. Clay Mathematics Institute. *"P vs. NP." http://www.claymath.org/millennium/P_vs_NP/.Cook, S. "The P versus NP Problem."* , [online] [cited 18.09.2021].
   Available: http://www.claymath.org/millennium/P_vs_NP/Official_Problem_Description.pdf.

6. Crescenzi, P. and Kann, V. (Eds.). *"A Compendium of NP Optimization Problems."* [online] [cited 05.09.2021]
   Available: http://www.nada.kth.se/~viggo/wwwcompendium/wwwcompendium.html.

7. Greenlaw, R.; Hoover, H. J.; and Ruzzo, W. L. *Limits to Parallel Computation: P-Completeness Theory.* Oxford, England: Oxford University Press, Published: 1995, [cited 01.11.2021].

8. Smale, S. *"Mathematical Problems for the Next Century."* Math. Intelligencer 20, No. 2, 7-15, Published: 1998, [cited 05.11.2021]

9. S. Smale, Ed. V. Arnold, M. Atiyah, P. Lax, and B. Mazur.*"Mathematical Problems for the Next Century." In Mathematics: Frontiers and Perspectives* 2000 Providence, RI: Amer. Math. Soc., Published: 2000. [cited 17.11.2021]

10. Rishab Prabhu. *Execution of an C, C++ code*. Published: 21.12.2018 [online] [accessed 16.10.2021]
    Available: https://www.geeksforgeeks.org/how-does-a-c-program-executes/