UDC 519.85
DOI https://doi.org/10.32782/IT/2024-2-13

Denys OLIANIN
Master of Faculty of Cybernetics and Computer Sciences, Taras Shevchenko National University of Kyiv, 60, Volodymyrska Str., Kyiv, Ukraine, 01601, denys.olianin@knu.ua
***ORCID:** 0009-0002-5266-3941*

To cite this article: Olianin, D. (2024). Optymizatsiia funktsii lohistychnykh vytrat na transportuvannia [Logistics transportation cost function optimization]. *Information Technology: Computer Science, Software Engineering and Cyber Security,* 2, 103–109, doi: https://doi.org/10.32782/IT/2024-2-13

LOGISTICS TRANSPORTATION COST FUNCTION OPTIMIZATION

Purpose of the study. *Optimizing transportation costs is a fundamental challenge in logistics management that requires advanced techniques to handle the complexities of geographical distances, transportation modes, and operational constraints. This article investigates the use of RMSProp, RMSProp with gradient clipping, and Proximal Gradient Descent methods in the optimization of transportation cost functions within logistics networks. By incorporating quality functions associated with both origin and delivery points, the study seeks to achieve a balance between cost reduction and the enhancement of service quality.*

Methodology*. We present a comparative analysis of the performance of these methods, focusing on their efficiency in terms of convergence rates and the quality of the solutions obtained. The research demonstrates that RMSProp and its variant with gradient clipping are particularly effective in navigating the solution space, offering fast convergence and high-quality solutions. Proximal Gradient Descent, on the other hand, shows promise in handling the discrete nature of logistics problems.*

Scientific novelty. *This study underscores the critical role of transportation cost optimization amidst the growing demands of global trade and logistics. As global trade expands, the need to minimize transportation expenses while maintaining service quality becomes increasingly vital. Advanced optimization techniques like RMSProp and Proximal Gradient Descent can lead to significant cost savings and improved operational efficiency, benefiting businesses involved in international trade.*

Conclusions. *In conclusion, this study highlights the effectiveness of RMSProp and Proximal Gradient Descent in optimizing transportation costs. It emphasizes the necessity of continuous innovation in logistics management to meet the evolving demands of global trade. Future research directions include the exploration of hybrid optimization techniques that combine the strengths of gradient-based methods, further enhancing the robustness and applicability of transportation cost optimization models in diverse logistical environments.*

Key words: *transportation cost optimization, Gradient descent methods, Logistics management, Evaluating functions, RMSProp, Proximal gradient descent.*

Денис ОЛЯНІН
магістр факультету кібернетики та комп'ютерних наук, Київський національний університет імені Тараса Шевченка, вул. Володимирська, 60, м. Київ, Україна, 01601
***ORCID:** 0009-0002-5266-3941*

Бібліографічний опис статті: Олянін, Д. (2024). Оптимізація функції логістичних витрат на транспортування. *Information Technology: Computer Science, Software Engineering and Cyber Security,* 2, 103–109, doi: https://doi.org/10.32782/IT/2024-2-13

ОПТИМІЗАЦІЯ ФУНКЦІЇ ЛОГІСТИЧНИХ ВИТРАТ НА ТРАНСПОРТУВАННЯ

Мета роботи*. Оптимізація транспортних витрат є фундаментальною проблемою в управлінні логістикою, яка потребує використання передових методів для вирішення складнощів, пов'язаних з географічними відстанями, видами транспорту та операційними обмеженнями. У цій статті досліджується використання методів RMSProp, RMSProp з обрізанням градієнта та Проксимального Градієнтного Спуску в оптимізації функцій транспортних витрат у логістичних мережах. Включаючи функції якості, пов'язані як з початковими, так і з кінцевими точками доставки, дослідження прагне досягти балансу між зниженням витрат та покращенням якості обслуговування.*

Методологія. *Представлено порівняльний аналіз ефективності цих методів, зосереджуючись на їх ефективності з точки зору швидкості збіжності та якості отриманих рішень. Дослідження демонструє, що RMSProp та його варіант з обрізанням градієнта є особливо ефективними у навігації в просторі*

рішень, забезпечуючи швидку збіжність та високоякісні рішення. З іншого боку, Проксимальний Градієнтний Спуск показує перспективи у вирішенні дискретних проблем логістики.

Наукова новизна. *Це дослідження підкреслює важливу роль оптимізації транспортних витрат у контексті зростаючих вимог глобальної торгівлі та логістики. З розширенням глобальної торгівлі необхідність зниження транспортних витрат, зберігаючи при цьому якість обслуговування, стає дедалі важливішою. Передові методи оптимізації, такі як RMSProp та Проксимальний Градієнтний Спуск, можуть призвести до значної економії витрат та покращення операційної ефективності, що приносить користь бізнесу, який займається міжнародною торгівлею.*

Висновки. *У висновку, це дослідження підкреслює ефективність RMSProp та Проксимального Градієнтного Спуску в оптимізації транспортних витрат. Воно наголошує на необхідності постійних інновацій в управлінні логістикою для задоволення еволюційних вимог глобальної торгівлі. Напрями майбутніх досліджень включають вивчення гібридних методів оптимізації, які поєднують сильні сторони методів, заснованих на градієнтах, ще більше підвищуючи надійність та застосовність моделей оптимізації транспортних витрат у різноманітних логістичних середовищах.*

Ключові слова: *оптимізація транспортних витрат, Методи градієнтного спуску, Управління логістикою, Оцінювальні функції, RMSProp, Проксимальний градієнтний спуск.*

Introduction. Optimizing transportation costs is a critical challenge in logistics management, essential for maintaining competitiveness and profitability in the global market. As businesses extend their operations across borders, they face complex factors such as geographical distances, diverse transportation modes, and various operational constraints. Traditional methods often fall short in addressing these complexities, highlighting the need for advanced optimization techniques.

Recent advancements in machine learning and optimization have introduced powerful algorithms capable of tackling these challenges effectively. This article explores the application of modern optimization algorithms–specifically RMSProp, RMSProp with gradient clipping, and Proximal Gradient Descent–in the context of transportation cost optimization within logistics networks. These methods, grounded in optimization theory and machine learning, offer efficient solutions by navigating the solution space to find cost-effective and high-quality transportation plans.

A key aspect of our study involves integrating quality functions related to both origin and delivery points, ensuring that cost reduction does not come at the expense of service quality. This balance is crucial for achieving optimal performance in logistics operations, as demonstrated by existing research on the topic (Crainic, Perboli, Rosano. 2018, p 410-418).

The research presented in this article includes a comparative analysis of RMSProp, its variant with gradient clipping, and Proximal Gradient Descent. We focus on their efficiency in terms of convergence rates and the quality of solutions obtained, providing valuable insights into their applicability in real-world logistics scenarios. Previous studies have shown the potential of gradient-based methods in various optimization problems (Duchi, Hazan, Singer, 2011, p 2121-2159; Kingma, Ba, 2014) which we aim to build upon in the context of logistics.

Moreover, this study underscores the importance of transportation cost optimization amidst the growing demands of global trade and logistics. With international trade volumes increasing, the need to minimize transportation expenses while maintaining service quality becomes more critical. Advanced optimization techniques like RMSProp and Proximal Gradient Descent can lead to significant cost savings and improved operational efficiency, benefiting businesses involved in global trade (Cordeau, Toth, Vigo, 2002, p 380-404; Laporte, 2007, p 811-819).

In conclusion, the findings of this research emphasize the necessity of continuous innovation in logistics management to meet the evolving demands of global commerce. By exploring and implementing advanced optimization techniques, logistics managers can ensure their operations remain both cost-effective and competitive in an increasingly complex and dynamic environment. This article contributes to the existing body of knowledge by demonstrating the practical benefits of modern optimization methods in transportation cost management, encouraging further research and application in this field.

Related works. The optimization of transportation costs is a well-explored area in logistics and operations research, encompassing a variety of methodologies and approaches. Traditional methods often involve linear programming, mixed-integer programming, and heuristic algorithms to minimize transportation costs while meeting service requirements. In recent years, advanced optimization techniques, particularly those leveraging gradient-based methods, have gained traction due to their ability to handle complex and high-dimensional problem spaces efficiently.

Gradient descent and its variants have been widely used for optimization problems in various domains. (Kingma, Ba, 2014) introduced Adam, an optimization method that computes

adaptive learning rates for each parameter, showing superior performance in training deep neural networks. Similarly, (Duchi, Hazan, Singer, 2011, p. 2121-2159) proposed Adagrad, an algorithm that adapts the learning rate based on the gradients, which has been effective in sparse data settings. Our approach leverages RMSProp and its variant with gradient clipping to handle the complexities of transportation cost optimization, providing a robust solution for point-to-point delivery scenarios.

Crainic (Crainic, Perboli, Rosano, 2018, p. 408-410) provided a comprehensive overview of intermodal freight transportation, highlighting the challenges and solutions in optimizing transportation networks. Their work emphasizes the importance of cost minimization while maintaining service quality. Toth and Vigo (Toth,Vigo 2014, p 463) explored vehicle routing problems, presenting various optimization models and algorithms tailored for transportation cost reduction. Our approach builds on these foundational works by introducing a differentiable quality function that enables the application of gradient-based optimization methods, thereby enhancing the efficiency and effectiveness of transportation cost management.

The primary advantage of our approach lies in the differentiability of the quality function, which allows for the application of gradient descent methods. This not only ensures faster convergence rates but also provides high-quality solutions. By incorporating both the quality functions of the origin and destination, along with the truck expenses and distance function, our model offers a comprehensive framework for transportation cost optimization.

Moreover, the use of RMSProp with gradient clipping helps to mitigate issues related to exploding gradients, ensuring stable and reliable optimization processes. Proximal Gradient Descent further enhances the model's capability to handle discrete and combinatorial aspects of logistics problems, providing a versatile and effective solution for real-world applications.

Our study contributes to the ongoing research in transportation cost optimization by introducing a novel approach that combines the strengths of gradient-based methods with a well-defined quality function. This approach not only addresses the complexities of point-to-point deliveries but also sets the stage for further advancements in the field, encouraging the exploration of hybrid optimization techniques that integrate gradient-based methods with heuristics and metaheuristics.

Methods of optimization. To address the complexities involved in optimizing transportation costs for point-to-point deliveries, we employ a set of advanced optimization algorithms. These algorithms are chosen for their robustness, efficiency, and ability to handle high-dimensional, non-linear optimization problems. In this section, we will describe the key algorithms used in our study: RMSProp, RMSProp with gradient clipping, and Proximal Gradient Descent. Each of these algorithms offers unique advantages and has been tailored to fit the specific requirements of our transportation cost function optimization model.

We suppose that our transportation cost function is convex, differentiable, and given by:

$$C(x_1, y_1, x_2, y_2, t) = \frac{Q(x_2, y_2)}{Q(x_1, y_2)} \cdot t \cdot D(x_1, y_1, x_2, y_2) \quad (1)$$

where x_1, y_1, x_2, y_2 – are the coordinates of the origin and delivery, t – truck cost coefficient, including expenses related to the driver and truck operation, $Q(x, y)$ – quality function of the location, $D(x_1, y_1, x_2, y_2)$ – distance function, in our research it is the function composition of Euclidian distances between two points. Therefore, we can use some gradient methods to minimize cost function.

We begin with an overview of the RMSProp algorithm, a popular choice in the field of machine learning for its adaptive learning rate capabilities. Next, we explore the variant of RMSProp with gradient clipping, which helps in managing the issues related to exploding gradients, ensuring more stable training processes. Finally, we delve into the Proximal Gradient Descent method, which is particularly effective for optimization problems involving constraints and regularization terms, making it well-suited for our discrete logistics scenarios.

By leveraging these algorithms, we aim to achieve a balance between minimizing transportation costs and maintaining high service quality, thereby enhancing the overall efficiency of logistics operations.

RMSProp, an adaptive learning rate optimization algorithm introduced by (Boyd, Parikh, 2013, p 130), offers significant advantages for optimizing complex transportation cost functions in logistics. This algorithm addresses the limitations of traditional gradient descent methods, particularly in non-convex and high-dimensional optimization problems.

In our context, RMSProp is particularly useful due to its ability to dynamically adjust the learning rate for each parameter, ensuring efficient convergence and stability. This is crucial when dealing with the coordinates of origin and delivery points. By maintaining a moving average of the squared gradients, RMSProp normalizes these gradients,

preventing the learning rate from becoming excessively small or large. This capability is essential for effectively navigating the complex and potentially unstable cost landscape in logistics optimization.

Given the non-linear nature of our objective function, which involves multiple variables such as geographic coordinates and cost coefficients, RMSProp's adaptive learning rate helps in achieving faster convergence and more stable solutions. This makes it a valuable tool for enhancing the efficiency and accuracy of our transportation cost function optimization efforts. Formally, given a parameter vector θ, in our case it is the vector of coordinates of an origin and a delivery, and the constant of the expenses connected to the truck and an objective function $f(\theta)$ to be minimized, RMSProp updates the parameters as follows:

1. Compute gradient $\nabla f(\theta_t)$ with respect to θ at iteration t.

2. Update the exponentially decaying average of squared gradients:

$$E\left[\nabla f(\theta_t)^2\right] = \rho E\left[\nabla f(\theta_{t-1})^2\right] + (1-\rho)\left(\nabla f(\theta_t)\right)^2 \quad (2)$$

where ρ is a decay rate set to 0.999 after tunning.

3. Update the parameter vector θ:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{E\left[\nabla f(\theta_t)^2\right] + \epsilon}} \cdot \nabla f(\theta_t) \quad (3)$$

Where η – is the global learning rate 1e-6 and ϵ is the small constant to avoid division by zero.

RMSProp with Gradient Clipping, an enhancement of the RMSProp optimization algorithm, effectively addresses the challenges posed by exploding gradients during the optimization process. This technique is particularly advantageous for complex problems like transportation cost optimization, where the objective function can exhibit steep gradients and significant variability.

In the context of our problem, RMSProp with Gradient Clipping improves the stability and convergence of the optimization process. Gradient clipping, a technique used to limit the size of the gradients, prevents them from becoming excessively large and causing instability. By incorporating gradient clipping into RMSProp, we ensure more stable and reliable updates during the optimization process. This results in more efficient and accurate convergence, making it a valuable approach for optimizing transportation cost functions in logistics. The algorithm is defined by the following steps:

1. Compute gradient $\nabla f(\theta_t)$ with respect to θ at iteration t.

2. Clip the gradient to a maximum norm δ:

$$\nabla f(\theta_t) = \frac{\nabla f(\theta_t)}{\max\left(1, \frac{\left\|\nabla f(\theta_t)\right\|}{\delta}\right)} \quad (4)$$

where $\nabla f(\theta_t)$ – is the L2 norm of the gradient and δ is the clipping threshold.

3. Update the exponentially decaying average of squared gradients:

$$E\left[\nabla f(\theta_t)^2\right] = \rho E\left[\nabla f(\theta_{t-1})^2\right] + (1-\rho)\left(\nabla f(\theta_t)\right)^2 \quad (5)$$

where ρ is a decay rate set to 0.999 after tunning.

4. Update the parameter vector θ:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{E\left[\nabla f(\theta_t)^2\right] + \epsilon}} \cdot \nabla f(\theta_t) \quad (6)$$

Proximal gradient descent is an optimization algorithm designed to handle problems involving non-smooth regularization terms, making it particularly useful for structured optimization problems commonly encountered in machine learning and signal processing introduced in work (Hinton, Srivastava, Swersky, 2012). This algorithm extends the standard gradient descent method by incorporating a proximal operator, effectively managing the non-differentiable components of the objective function.

In our transportation cost optimization problem, we incorporate L2 regularization to separate the function into potentially non-differentiable and differentiable parts. The objective function in our optimization problem can be expressed as:

$$F(x) = f(x) + g(x) \quad (7)$$

where $f(x)$ is a smooth, differentiable function representing the primary cost function, and $g(x)$ is a non-differentiable function representing the regularization term.

In the context of L2 regularization, the objective function becomes:

$$F(x) = f(x) + \frac{\lambda}{2} \| x \|^2 \quad (8)$$

where $\frac{\lambda}{2} \| x \|^2$ is the L2 regularization term, λ is the regularization parameter, and $\| x \|^2$ denotes the squared Euclidean norm of the parameter vector x.

The Proximal Gradient Descent algorithm alternates between a gradient descent step on the smooth part $f(x)$ and a proximal operator step on the non-differentiable part $g(x)$

1. Start with an initial guess x_0.

2. Update the parameters by performing a gradient descent step on the smooth part:

$$y_{k+1} = x_k - \eta \nabla f(x_k) \quad (9)$$

where η is the learning rate, and $\nabla f(x_k)$ is the gradient of f at x_k.

3. Apply the proximal operator associated with the L2 regularization term:

$$x_{k+1} = prox_{\eta\lambda}(y_{k+1}) \quad (10)$$

The proximal operator for L2 regularization is defined as:

$$prox_{\eta\lambda}(v) = \frac{v}{1+\eta\lambda} \quad (11)$$

This step ensures that the parameters are regularized, promoting smaller values and preventing overfitting.

4. Repeat the gradient descent and proximal operator steps until convergence or for a predetermined number of iterations.

Experimental results. In this section, we evaluate the performance of the proposed algorithms by comparing them. All code is written in Python using the NumPy, Matplotlib, and SymPy libraries. The experiments are conducted on a MacOS laptop with an M1 Pro CPU and 16 GB RAM.

The quality of a logistics area depends on several factors, including the number of nearby facilities such as warehouses, hubs, airports, and docks, as well as the presence of production lines, commercial stores, and medical facilities. Generally, urban areas have the highest quality, while rural areas have the lowest. Our quality function is defined as follows:

$$\begin{aligned} Q(x,y) = {} & 3(1-x)^2 \cdot e^{-x^2-(y+1)^2} - 10\left(\frac{x}{5} - x^3 - y^5\right) \cdot e^{-x^2-y^2} - \\ & - \frac{1}{3} \cdot e^{-(x+1)^2-y^2} + 5\left(\frac{x}{5} - x^5 - y^7\right) \cdot e^{-x^2-y^2} - \\ & - \frac{1}{3} \cdot e^{-(x+1)^2-y^2} + 2\left(\frac{x}{5} - x^5 - y^7\right) \cdot e^{-x^4-y^4} - \\ & - \frac{1}{3} \cdot e^{-(x+4)^2-y^2} + 10 \end{aligned} \quad (12)$$

The surface plot of the quality function is on the figure 1, perfectly represents the nature of the logistics areas.

In our experiment distance function shortly can be expressed as:

$$D(x_1, y_1, x_2, y_2) = d(x_1, y_1, x_2, y_2) + \frac{k}{d(x_1, y_1, x_2, y_2)},$$

where x_1, y_1- are coordinates of an origin, x_2, y_2 – coordinates of a destination, k=50 is coefficient, that is needed to avoid small transportation cost to close objects, as we always have minimum fair for transportation, $d(x_1, y_1, x_2, y_2)$ – Euclidian distance between two points.

Therefore, our transportation cost function can be expressed as:

$$C(x_1, y_1, x_2, y_2, t) = \frac{Q(x_2, y_2)}{Q(x_1, y_1)} \cdot t \cdot D(x_1, y_1, x_2, y_2)$$

where t is a coefficient of the cost of truck type in our experimental case for the sake of simplicity is 1,2.

The provided graph on figure 2 and data in table 1 illustrate the performance of three optimization methods–RMSProp, RMSProp with Gradient Clipping, and Proximal Gradient Descent–on the transportation cost function.

The convergence speed of these methods varies significantly. RMSProp converged in 1045 iterations, taking approximately 0.7325 ms per iteration. RMSProp with Gradient Clipping achieved the fastest convergence among the methods, requiring only 421 iterations, with a time of 0.6968 ms per iteration. Proximal Gradient Descent converged in 617 iterations, with the lowest time per iteration at 0.6832 ms.

The cost function values indicate distinct behaviors. All methods show a steep decline in the cost function value within the first few hundred iterations, indicating rapid initial convergence. However, RMSProp and RMSProp with Gradient Clipping stabilize at a higher cost function value compared to the Proximal Gradient Descent, which suggests that while they converge quickly, the final optimized value is not as low as that achieved by the Proximal method. Proximal Gradient Descent reaches a lower cost function value, indicating better overall optimization performance in minimizing the cost function.

In terms of stability, RMSProp and RMSProp with Gradient Clipping show some variability in the early stages but stabilize quickly, with gradient clipping helping RMSProp achieve faster stabilization. Proximal Gradient Descent demonstrates the most stable and consistent decrease in the cost function without significant fluctuations.

Conclusions. In conclusion, our research highlights the effectiveness of gradient-based optimization methods for addressing transportation cost functions, especially when the functions are differentiable. Among these methods, the Proximal gradient descent method stands out, demonstrating superior performance in terms of convergence speed and solution quality. However, it is important to recognize the challenges posed by local minima and steep gradients, particularly in the cases of RMSProp and RMSProp with gradient clipping. The comparative analysis of RMSProp, RMSProp with Gradient Clipping, and Proximal Gradient Descent reveals distinct strengths and

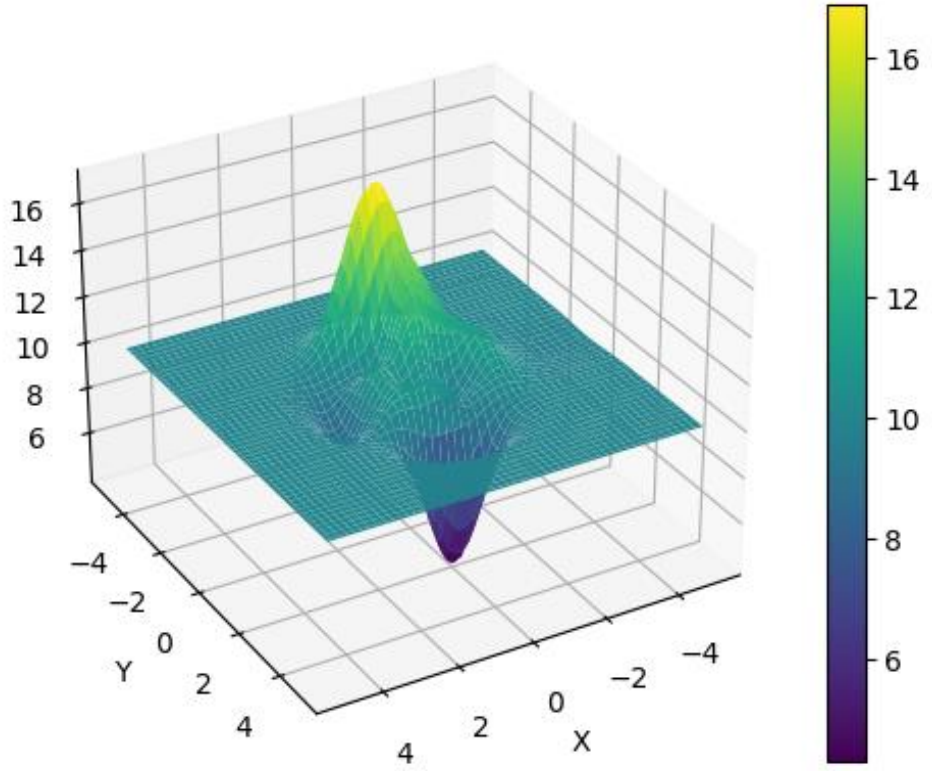

Fig. 1. Quality function surface plot

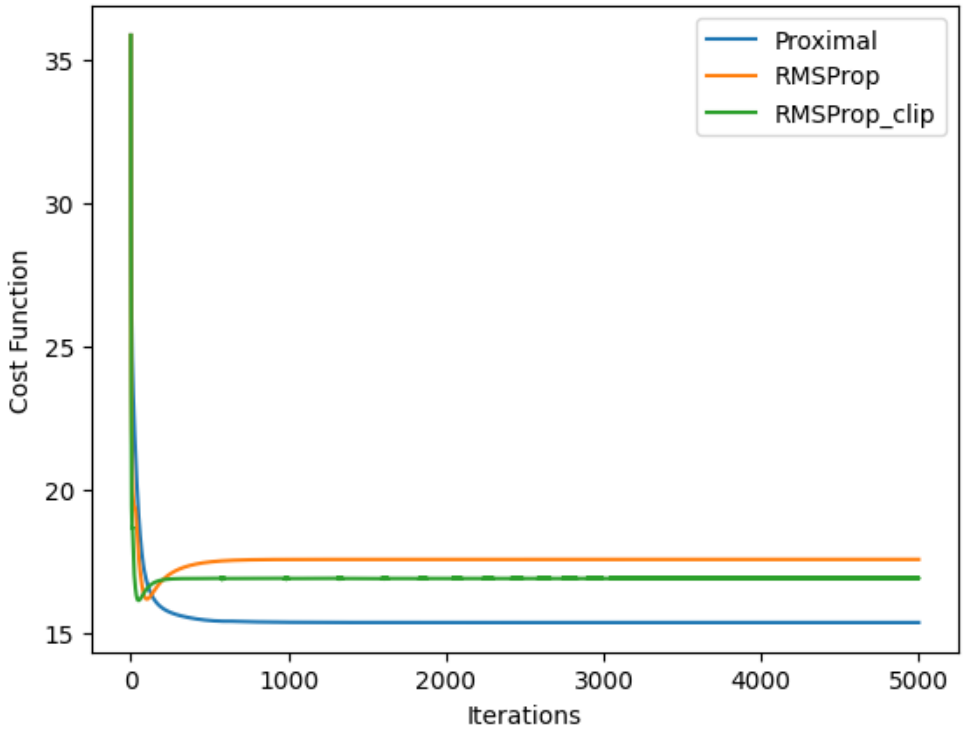

Fig. 2. Transportation cost function graph per iteration

Table 1

Method	Time per Iteration (ms)	Iterations to Converge
RMSProp	0.7325	1045
RMSProp with Gradient Clipping	0.6968	421
Proximal Gradient Descent	0.6832	617

areas of suitability for each method. RMSProp with Gradient Clipping exhibits the fastest convergence in terms of iterations and provides a stable optimization process due to gradient clipping, making it best for scenarios where rapid convergence is essential, and slight deviations in the final cost function value can be tolerated. Proximal Gradient Descent achieves the lowest cost function value, indicating superior optimization performance. It combines stability and efficiency, taking the least time per iteration, making it ideal for optimization problems where achieving the absolute minimum cost function value is crucial, even if it requires slightly more iterations than the fastest converging method. RMSProp provides a balance between convergence speed and optimization quality, though not excelling in either aspect compared to the other two methods and is suitable for general use when both convergence speed and the quality of the optimized result are important, but not critical.

Overall, the choice of optimization method should be guided by the specific requirements of the problem at hand. For rapid convergence, RMSProp with Gradient Clipping is optimal, whereas for achieving the best minimized cost function value, Proximal Gradient Descent is preferred.

Future research could benefit from expanding the quality function and truck function to reflect more realistic scenarios, incorporating real geographic locations and fuel costs.

The significance and potential of optimizing transportation cost functions remain clear, particularly with recent advancements in algorithmic techniques. As the field progresses, integrating innovative methodologies and refining existing approaches will be essential for improving the efficiency of transportation cost optimization strategies.

BIBLIOGRAPHY:

1. Boyd S., Parikh N. Proximal Algorithms. Now Publishers, 2013. 130 p.

2. Cordeau J.-F., Toth P., Vigo D. A Survey of Optimization Models for Train Routing and Scheduling. *Transportation Science*. 1998. Vol. 32, no. 4. P. 380–404. URL: https://doi.org/10.1287/trsc.32.4.380 (date of access: 27.05.2024).

3. Crainic T. G., Perboli G., Rosano M. Simulation of intermodal freight transportation systems: a taxonomy. *European Journal of Operational Research*. 2018. Vol. 270, no. 2. P. 401–418. URL: https://doi.org/10.1016/j.ejor.2017.11.061 (date of access: 27.05.2024).

4. Duchi J., Hazan E., Singer Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*. 2011. Vol. 12, no. 61. P. 2121–2159.

5. Hinton G., Srivastava N., Swersky K. Neural networks for machine learning, Lecture 6a: Overview of mini-batch gradient descent. *CSC321 – Introduction to Neural Networks and Machine Learning*. URL: http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf (date of access: 27.05.2024).

6. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization. arXiv, 2014. 15 p. (Preprint. arXiv:1412.6980). URL: https://doi.org/10.48550/arXiv.1412.6980 (date of access: 27.05.2024).

7. Laporte G. What you should know about the vehicle routing problem. *Naval Research Logistics*. 2007. Vol. 54, no. 8. P. 811–819. URL: https://doi.org/10.1002/nav.20261 (date of access: 27.05.2024).

8. Toth P., Vigo D. Vehicle routing: Problems, methods, and applications. Philadelphia : Society for Industrial and Applied Mathematics, 2014. 463 p.

REFERENCES:

1. Boyd, S. & Parikh, N. (2013). Proximal Algorithms. Now Publishers, 130 p.

2. Cordeau, J. F., Toth, P. & Vigo, D. (2002). A survey of optimization models for train routing and scheduling. *Transportation Science*, 32(4), 380–404

3. Crainic, T. G., Perboli, G. & Rosano, M. (2018). Simulation of intermodal freight transportation systems: A taxonomy. *European Journal of Operational Research*, 220(3), 410–418.

4. Duchi, J., Hazan, E. & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul), 2121–2159.

5. Hinton, G., Srivastava, N. & Swersky, K. (2012). Neural networks for machine learning, Lecture 6a: Overview of mini-batch gradient descent. *CSC321 – Introduction to Neural Networks and Machine Learning.* Retrieved from http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.

6. Kingma, D. P. & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

7. Laporte, G. (2007). What you should know about the vehicle routing problem. *Naval Research Logistics*, 54(8), 811–819.

8. Toth, P. & Vigo, D. (2014). Vehicle Routing: Problems, Methods, and Applications. SIAM Monographs on Discrete Mathematics and Applications.