артиллерийское управление, 3 центральный научно-исследовательский институт.

8. Савельєв К. К. Основи будови і конструкції артилерійських установок // К. К. Савельєв. Севастополь: СВМІ, 2007.
9. Савельєв К. К. Основи будови і конструкції артилерійських установок. Альбом схем і рисунків. // К. К. Савельєв. Севастополь: СВМІ, 2007.
10. Артиллерия и ракеты // Воениздат, 1968.
11. Стрільба артилерії // В. М. Петренко, В. Є. Житник, В. І. Макеєв та ін. Суми: Сумський державний університет, 2012. 760 с.

Roman Dorosh

Vasyl Stefanyk Precarpathian National University,

Ivano-Frankivsk, Ukraine

roman.dorosh.17@pnu.edu.ua

IMPROVING AND IMPLEMENTING BLOCKCHAIN-BASED SYSTEM IN EDUCATIONAL FIELD

This paper proposed new unified algorithm for plagiarism detection based on Ethereum blockhain ecosystem with similarity detection SimHash algorithm which works on around multi - billion pages in a reasonable amount of time. Blockchain will act as a transparent technology against unauthorized access through a system of social and smart contracts.

Keywords: *similarity detection; simhash; blockchain; smart-contracts; education; natural language processing; ethereum virtual machine.*

Blockchain also provides the main characteristics: autonomy, decentralization, fault-tolerance, distribution, open, immutability and transparency [1], which led to interest of researchers into this area.

Blockchain started to gain its popularity after the publication of Satoshi Nakamoto original paper [2], where the main goal was to show the solution to achieve consensus through decentralization of servers. So, the “mainers” get rewards through supporting the infrastructure by serving the blockchain on their own hardware or deploying to cloud-based solutions. This kind of algorithm is called Pow (Proof of Work), which forms transactions which cannot be changed without redoing the PoW.

After success of the Bitcoin and realizing the limitation of this approach to build blockchain, the second popular system operating on PoS (Proof of Stake) called Ethereum [3] appears. Ethereum has it’s own ecosystem with support to build decentralized application through API using smart-contracts as rule-based systems. It gives developers power to build web application and deploy it to Ethereum blockchain.

Common approach to detect plagiarism is to apply natural language processing to extract useful data and normalize it; after that the simple cosine similarity, slightly changed version of cosine similarity [4] or TF-IDF algorithm can be used to find similarity score between selected documents.

Machine Learning and particularly Neural Networks can be applied to such areas as software defined demodulation of weak radio signals, which proposed in [5, 6] papers and efficient identity document detection [7] etc. Research [8] shows the importance of introducing new ways to improve education, especially online engineering education.

Pungilă et al [9] proposed to use Blockchain-Based Non-Fungible Tokens approach with different similarity measurements. They conclude that their approach can be use in the real-world scenarios and has significant speed-up improvement to classic similarity measurements. However, there is lots of drawbacks: large amounts of NFT (converts into large amount of costs), cannot detect the slightly altered version of the same data which presented differently (using the SHA256 hash).

Main advantages of SimHash algorithm: scalability and speed. The original paper [10] shows practically usefulness for identifying near-duplicates in documents belonging to large file storage (multi-billion repository).

So, based on implementation areas and attributes of SimHash, this algorithm was chosen as engine for plagiarism detection. In this approach, we look through set of documents $S = \{D1, D2, D3, \ldots Dn\}$ and find SimHash of every entry $SHd = SimHash(D)$. After that, every entry hash SHd compares to all documents S one by one by corresponding distance between them. The result of this operation will be number between [0, 1] which can be translated to corresponding percentage. Also, set of hashes Sh stored on the Ethereum blockchain through smart contract and cannot be deleted by central authority. SimHash algorithm gives power to cluster similar documents into groups and detect plagiarism of many entries by one query to the database. The main disadvantage here that algorithm cannot show the accurate places where the cheater copies from the original paper.

The text processing uses the following steps: tokenization, removing stop words, stemming, lemmatization and conversion to word embeddings to support multi-language plagiarism detection.

In this paper I propose the unification of existing plagiarism detection methods, i.e. SimHash and decentralized blockchain technology into one algorithm (Figure 1).

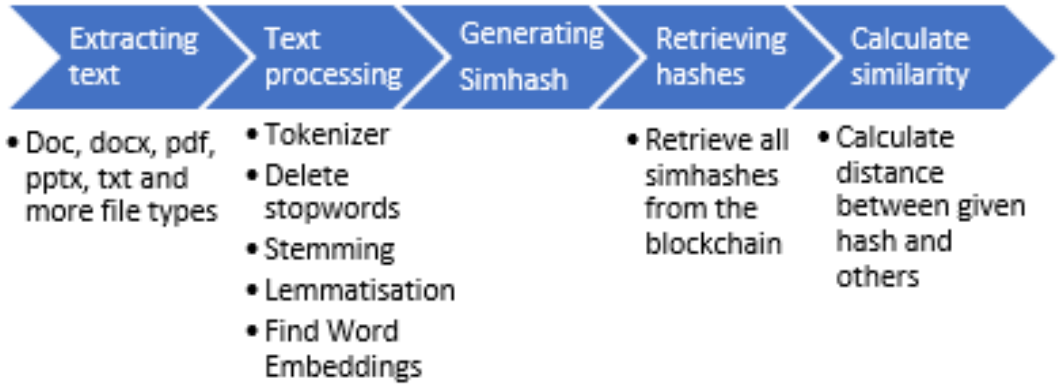

Figure 1 – Unified Algorithm

Finding plagiarism and calculate the similarity is not the easy task, especially when it comes to the multi-language comparison. For evaluation of the algorithm performance and accuracy I've implemented the next pseudo-algorithm (Figure 2).

Algorithm 2 Evaluation similarity pseudo algorithm

for `each added` *Document* **do**
  `process` *Document*
  `generate` *SimHash* `for processed` *Document*
  if *generatedSimHash* can be found on the Blockchain **then**
    added *Document* totally plagiarized ($S \leftarrow 1$)
  else
    calculate similarity S
  end if
return 10 most common *SimhHashes* and their S (similarity) score in range [0, 1]
end for

Figure 2 - Evaluation similarity pseudo algorithm

The dataset [11] for evaluation proposed in this paper consists of 143,000 news articles from 15 major publications but used only 10,000. The language of the articles is English, because multi-language support is not fully implemented. Also, I can formalize this task, as a classification with two classes (plagiarism - 0 and not plagiarism - 1).

I have performed testing on Processor 11th Gen Intel(R) Core(TM) i7-1185G7 @ 3.00GHz, 2995 Mhz, 4 Core(s), 8 Logical Processor(s) with 32 GB RAM and 512 GB SSD drive, OS is Windows 10. The Python library for visualizing metric results *sklearn* was used.

So, the results of testing my algorithm is given below on the Figure 3 and Figure 4 (ROC AUC score).

	precision	recall	f1-score	support
0	0.90	0.90	0.90	4988
1	0.90	0.91	0.90	5012
accuracy			0.90	10000
macro avg	0.90	0.90	0.90	10000
weighted avg	0.90	0.90	0.90	10000

Figure 3 – Proposed algorithm metrics

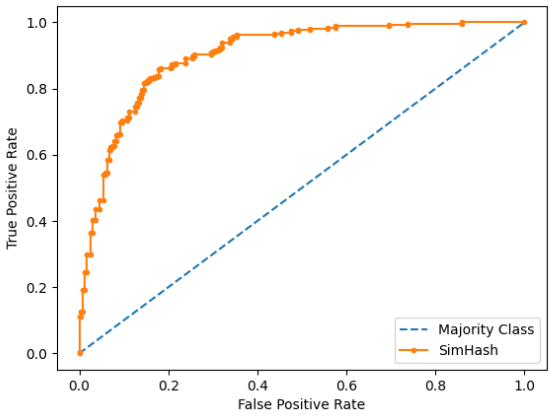

Figure 4 – ROC AUC score

After comparing new algorithm to TF-IDF algorithm the overall performance was better by 7%.

The proposed unified algorithm operates well on detecting exact or partial copies of the documents; also, the performance doesn't downgrade with scalability and size of the documents; changing words to their synonyms or paraphrasing can be detected too; identification of the multi-language plagiarism performs on the basic level, to achieve more precise solution, the complex language analysis must be conducted. Also, the SimHash algorithm operates quite good on the large number of documents (around billions), so the speed of plagiarism detection is stable one low and high document number.

REFERENCES

1. M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli and M. H. Rehmani, "Applications of Blockchains in the Internet of Things: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1676-1717, Secondquarter 2019, doi: 10.1109/COMST.2018.2886932.
2. S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", Aug 2022, [online] Available: https://bitcoin.org/bitcoin.pdf.
3. V.Buterin, "Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform", Aug 2022, [online] Available:

https://ethereum.org/669c9e2e2027310b6b3cdce6e1c52962/Ethereum_Whitepaper_-_Buterin_2014.pdf.

4. Anzelmi, Daniele, Domenico Carlone, Fabio Rizzello, Robert Thomsen and Dil Muhammad Akbar Hussain. “Plagiarism Detection Based on SCAM Algorithm.” *IMECS 2011* (2011).
5. M. Kozlenko, I. Lazarovych, V. Tkachuk and V. Vialkova, "Software Demodulation of Weak Radio Signals using Convolutional Neural Network," *2020 IEEE 7th International Conference on Energy Smart Systems (ESS)*, 2020, pp. 339-342, doi: 10.1109/ESS50319.2020.9160035.
6. M. Kozlenko and V. Vialkova, "Software Defined Demodulation of Multiple Frequency Shift Keying with Dense Neural Network for Weak Signal Communications," *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2020, pp. 590-595, doi: 10.1109/TCSET49122.2020.235501.
7. M. Kozlenko,V. Sendetskyi, O. Simkiv, N. Savchenko, A. Bosyi, “Identity Documents Recognition and Detection using Semantic Segmentation with Convolutional Neural Network (short paper)”. *Cybersecurity Providing in Information and Telecommunication Systems 2021* (CPITS), 2021, pp. 234-242.
8. Dutchak, M., Kozlenko, M., Lazarovych, I., Lazarovych, N., Pikuliak, M., Savka, I. “Methods and Software Tools for Automated Synthesis of Adaptive Learning Trajectory in Intelligent Online Learning Management Systems”. *Innovations in Smart Cities Applications Volume 4. SCA 2020. Lecture Notes in Networks and Systems*, vol 183. Springer, Cham. https://doi.org/10.1007/978-3-030-66840-2_16.
9. C. Pungilă, D. Galis, V. Negru, “A New High-Performance Approach to Approximate Pattern-Matching for Plagiarism Detection in Blockchain-Based Non-Fungible Tokens (NFTs)”, arXiv:2205.14492 [cs.CR], May 2022.
10. G.S. Manku, A. Jain, A.D. Sarma, “Detecting Near-Duplicates for Web Crawling”, research.google.com,

https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/33026.pdf (accessed Aug. 23, 2022).

11. "All the news dataset". Kaggle.com https://www.kaggle.com/datasets/snapcrack/all-the-news (accessed Aug. 27, 2022).

Hajiyeva N.S., Rashidov A.M.

Institute of Applied Mathematics of Baku State University,

Baku, Azerbaijan

n.hajiyeva88@gmail.com, anar.reshidov1993@gmail.com

DETERMINATION OF OPTIMAL REGULATOR OF OSCILLATORY SYSTEMS WITH LIQUID DAMPERS DURING OIL PRODUCTION BY SUCKER ROD PUMPING UNIT

Abstract. In the paper the problem of determining optimal regulator of oscillatory systems with liquid dampers is considered.

Keywords: Oscillating system, fractional derivative, Euler-Lagrange method.

Let the motion of oscillatory systems with liquid dampers is described by the system of ordinary linear differential equations with fractional derivatives and boundary periodic conditions during oil production by sucker rod pumping unit (Fig. 1) as follows [2, 6, 7]

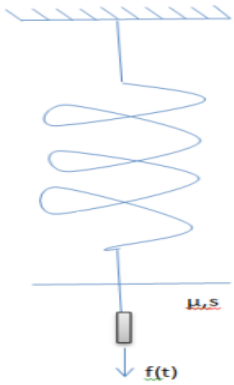

Figure 1. – Schematic representation of the problem.