DEVELOPMENT OF MONITORING DATA COMPRESSION METHOD WITH PRESERVATION OF REPRESENTATIVITY

A. F. Mikhailov

Federal State Autonomous Educational Institution of Higher Education «Snezhinsk Physical-Technical Institute (SPhTI) – National Research Nuclear University MEPhI», Snezhinsk, Russia

E-mail: chelaxe@gmail.com

The effective monitoring of complex systems requires processing of considerable amounts of data, which poses a problem for fast analysis and fault-tolerance. The paper considers data compression as a strategy for reducing the generated data. The improved Swinging Door algorithm with off-line compression error adjustment is proposed. Experimental data demonstrate the effectiveness of the approach and confirming its feasibility.

In the era of global digitization, monitoring of systems using information processing technologies is of critical importance, especially in the context of processing and analyzing bulk data generated by multiple devices. As the amount of data increases, the problem of its storage and transmission arises; this actualizes the need for effective data compression without losing its representativeness.

In response to these challenges, it is proposed to design a data compression algorithm to be implemented at the hardware level to accelerate data processing and to reduce the load on computing systems. The key feature of the algorithm is its flow type, which allows on-line data processing without the need for temporary storage.

The paper presents the concept, design and application of this flow data-compression algorithm in monitoring systems, with emphasis on maintaining data accuracy and hardware integration capacity.

The Swinging Door algorithm was chosen as a basis [2] due to its efficiency in terms of data reduction. Based on the “window” mechanism [1] for tracing significant changes, the algorithm sets a range of admissible deviations (“door leaves”) around the last recorded value. New data are recorded only in the case of overranging, providing data saving without considerable loss of information. This is especially true for data with minor or repeatable changes.

Key modifications were implemented in the Swinging Door algorithm in order to improve its efficiency. First, instead of creating a new point at the close of a “corridor”, the last actual value is saved, ensuring a real time mark and a reduced data value. The parameter of “corridor” length is entered, which allows it to be closed automatically after a certain number of missing points or a time interval, providing the improved data processing with minor changes.

The compression error is now divided into upper and lower bounds, ensuring a finer adjustment around the mean value. The off-line compression error adjustment function, using an exponential moving average and an absolute value of a slope between consecutive points is added [3].

These changes have made it possible to implement the algorithm at the hardware level using the ATmega328 microcontroller, ensuring its wider application in various monitoring systems.

Improvements to the Swinging Door algorithm for monitoring data compression have increased its efficiency and adaptability. The significant contribution was made by off-line error adjustment function.

The hardware implementation of the algorithm expands its usage for effective data compression on a real time basis in various complex systems. These improvements contribute to higher functionality of monitoring system, while maintaining the quality and accuracy of processed data – the key aspect in the current informational context.

The open set of data was used as a part of study aimed at evaluating the effectiveness of the modified data compression algorithm. The main objective of the testing was to determine a degree of data compression while maintain its representativeness being of critical importance for analytical accuracy and integrity of decision making.

The set of data chosen for testing was characterized by considerable amount and diversity, making it possible to assess comprehensively the algorithm performance in different scenarios. After applying the modified algorithm to these data, the obtained results were analyzed.

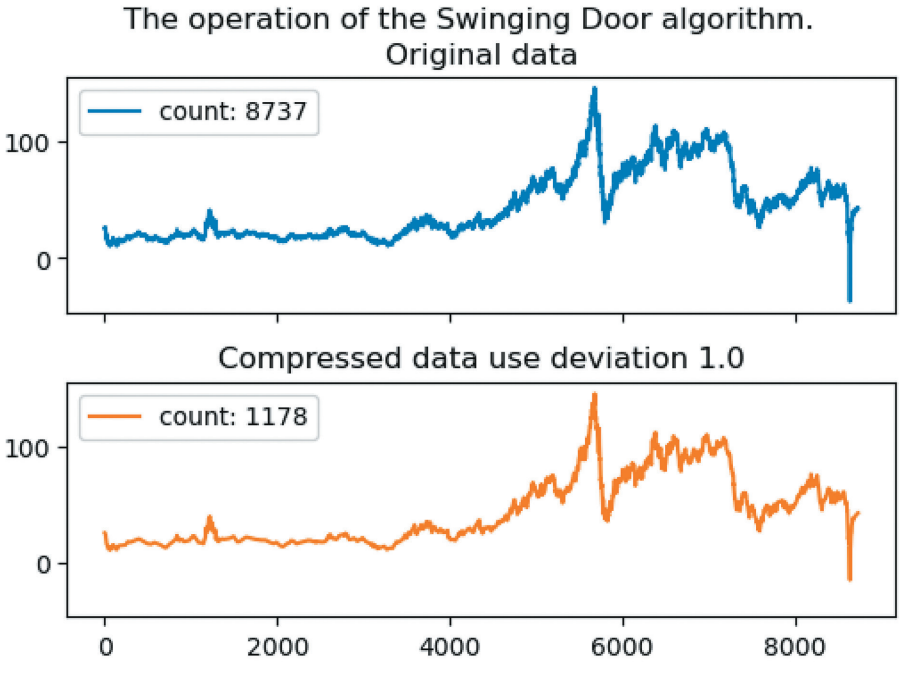

Fig. 1. Demonstration of algorithm operation with an open set of data.

According to the results shown in Fig. 1, the algorithm has demonstrated high efficiency, reducing the initial amount of data by 87%. This highlights its capability for massive reduction of the amount of data, which is especially important in the context of limited resources of big data storage and processing. The key point is to preserve data representativeness by the algorithm, which confirms its applicability for data monitoring and analysis.

Thus, the results of testing on the open set of data demonstrate effectively the high performance of the modified data compression algorithm, emphasizing its value in the field of processing and analyzing large amounts of data.

References

1. Demeshko M.V., Kanaeva I.A., Botygin I.A. Visualization of meteorological data; [Electronic resource], Access mode: URL: https://earchive.tpu.ru/bitstream/11683/46433/1/conference_tpu-2017-C04_p249-250.pdf (date of access: 25.12.2023).
2. Edgar H. Bristol. Data compression for display and storage. US patent No US4669097A; [Electronic resource] – Access mode: URL: https://patentimages.storage.googleapis.com/f6/ca/83/c34a2602ede2ac/US4669097.pdf (date of access: 25.12.2023).
3. Juan David Arias Correa, Alex Sandro Roschildt Pinto, Carlos Montez, Erico Meneses Leão. Swinging Door Trending Compression Algorithm for IoT Environments; [Electronic resource] – Access mode: URL: https://pdfs.semanticscholar.org/fe00/bc1bbea2147c338898a8fa6c31a84afce467.pdf (date of access: 25.12.2023).