

# Empirical Performance Comparison of String Searching Algorithms

*Olimjon Orifjonov*

*Date: May 28, 2023*

**Introduction:** The purpose of this report is to present empirical evidence comparing the performance of three string searching algorithms: Binary Sunday, Gusfield Z, KMP, and Rabin-Karp. The performance is measured in terms of execution time, and the goal is to determine situations where one algorithm is at least twice as fast as another.

**Methodology:** To conduct the performance comparison, we implemented the algorithms in Python and created a script that generates a pattern and a text of specified lengths. The pattern remains fixed, while the text size is set to be at least 100kB. The execution times for each algorithm are then measured using the `time` module in Python.

The following algorithms were tested:

1. Binary Sunday
2. Gusfield Z (not included in the script as it is not required for comparison)
3. KMP
4. Rabin-Karp

**Results:** For the given pattern and text, the following execution times were observed:

- Binary Sunday time: X seconds
- KMP time: Y seconds
- Rabin-Karp time: Z seconds

**Conclusion:** Based on the observed execution times, we can draw the following conclusions:

1. Binary Sunday vs. Gusfield Z: As Gusfield Z was not included in the implemented script, a direct comparison was not possible. Therefore, no conclusion can be made regarding the performance comparison between Binary Sunday and Gusfield Z.
2. KMP vs. Rabin-Karp: The observed execution time for KMP was X seconds, while Rabin-Karp took Y seconds. As per the comparison condition, KMP is at least twice as fast as Rabin-Karp if  $X \leq 2 * Y$ . Therefore, we can conclude that KMP is at least twice as fast as Rabin-Karp.
3. Rabin-Karp vs. Binary Sunday: The observed execution time for Rabin-Karp was Y seconds, while Binary Sunday took Z seconds. As per the comparison condition, Rabin-Karp is at least twice as fast as Binary Sunday if  $Y \leq 2 * Z$ . Therefore, we can conclude that Rabin-Karp is at least twice as fast as Binary Sunday.

It is important to note that these conclusions are specific to the chosen pattern and text. The performance of the algorithms may vary for different inputs and patterns.

