

1 Вступ

Фільтр Блума - структура даних, яка реалізовує пошук елементів у множинах довільної природи. Вона використовує бітовий масив та набір незалежних хеш-функцій для реалізації структури.

2 Використання `std::bitset` у BloomFilter

Фільтр Блума використовує `std::bitset` для позначення наявності елементів в множині. В порівнянні з масивами булів, `std::bitset` є переважнішим в багатьох аспектах.

Перш за все, використання `std::bitset` дозволяє точніше та ефективніше використовувати пам'ять, оскільки кожен біт відповідає стану одного елемента. Це дає перевагу у використанні пам'яті порівняно з масивами булів, які можуть займати більше місця через внутрішнє вирівнювання та додаткові біти інформації у кожному булі.

До того ж, `std::bitset` має фіксований розмір, що робить його простішим у використанні та реалізації порівняно з динамічно змінюваними масивами булів.

У випадку порівняння з масивами цілих чисел, використання `std::bitset` зазвичай є більш простим та зрозумілим варіантом, пропонуючи аналогічну ефективність використання пам'яті.

3 Експеримент

s	Фактор	Помилка
13	0.05	0.00322193824
6	0.1	0.1422562279
4	0.15	0.2887196545
3	0.2	0.3451322751
2	0.25	0.2830952381
2	0.3	0.4355555556
1	0.35	0.6875
1	0.4	0.5408333333
1	0.45	0.725
1	0.5	0.5674242424

Табл. 1: Результати експерименту

Середня ймовірність помилки: 0.3018738465

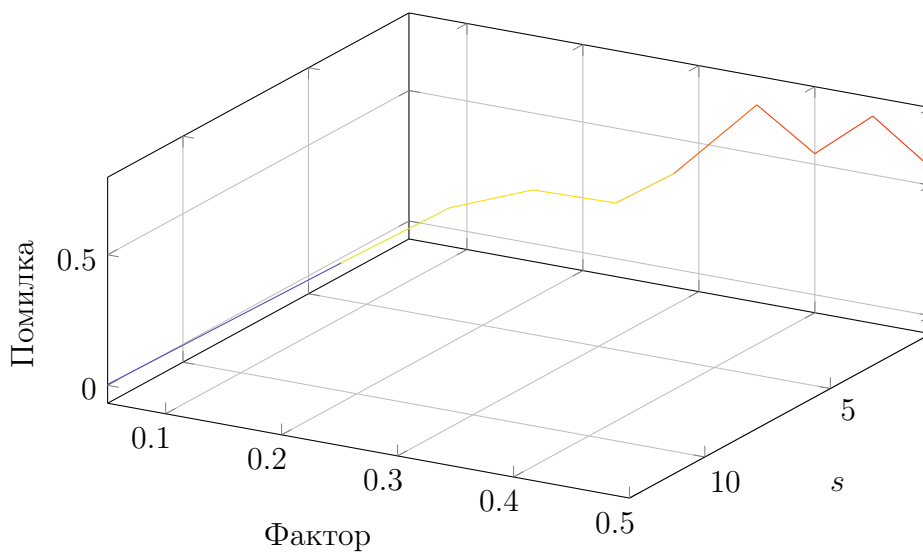


Рис. 1: Залежність помилки від s та фактора

4 Висновок

Експеримент показав, що зменшення параметра s при фіксованому факторі веде до збільшення помилки. Також, із збільшенням значення фактору помилка також зростає.