

# Comprehensive Benchmark Report: File Search Algorithms

Generated: 2025-11-03 06:56:31  
Tested file sizes: 10000, 50000, 100000, 250000, 500000, 1000000

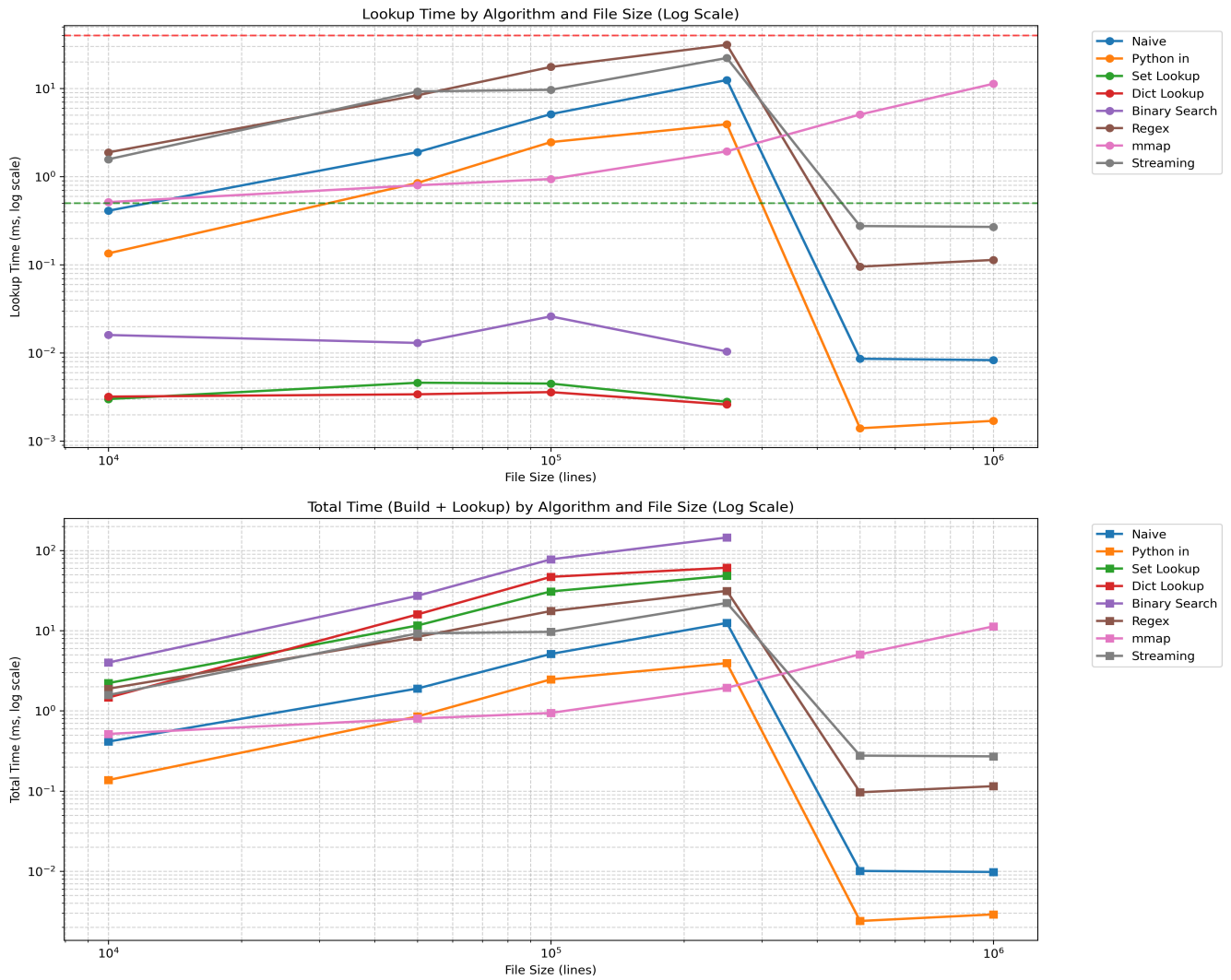
## Executive Summary

**Recommended algorithms (meet 0.5ms requirement):** Dict Lookup, Set Lookup, Binary Search

## Performance Requirement Validation (250k lines)

Algorithm	Scenario	Avg Time (ms)	Requirement	Status
Naive	REREAD_ON_QUERY=True	5.635	40ms	PASS
Python in	REREAD_ON_QUERY=True	2.007	40ms	PASS
Set Lookup	REREAD_ON_QUERY=False	0.005	0.5ms	PASS
Binary Search	REREAD_ON_QUERY=False	0.016	0.5ms	PASS
mmap	REREAD_ON_QUERY=True	1.624	40ms	PASS
Streaming	REREAD_ON_QUERY=True	27.990	40ms	PASS

## Performance Analysis Charts



## Algorithm Performance Summary

Algorithm	Avg Lookup (ms)	Max File Size	Meets 0.5ms	Scalability
Naive	3.319	1,000,000	NO	Good
Python in	1.230	1,000,000	NO	Good
Set Lookup	0.004	250,000	YES	N/A
Dict Lookup	0.003	250,000	YES	N/A
Binary Search	0.016	250,000	YES	N/A
Regex	9.885	1,000,000	NO	Good
mmap	3.425	1,000,000	NO	Moderate
Streaming	7.197	1,000,000	NO	Good

## Detailed Results (Sample)

Algorithm	File Size	Found	Build Time (ms)	Lookup Time (ms)	Total Time (ms)
Naive	10000	1	0.0033	0.4105	0.4138
Python in	10000	1	0.0025	0.1348	0.1373
Set Lookup	10000	1	2.2159	0.003	2.2189
Dict Lookup	10000	1	1.4625	0.0032	1.4657
Binary Search	10000	1	3.9859	0.016	4.0019
Regex	10000	1	0.0044	1.8842	1.8886
mmap	10000	1	0.002	0.5139	0.5159
Streaming	10000	1	0.0026	1.5725	1.5751
Naive	50000	1	0.0037	1.8978	1.9015
Python in	50000	1	0.0032	0.8494	0.8526
Set Lookup	50000	1	11.5984	0.0046	11.603
Dict Lookup	50000	1	15.9474	0.0034	15.9508
Binary Search	50000	1	27.1867	0.013	27.1997
Regex	50000	1	0.0049	8.3646	8.3695
mmap	50000	1	0.0029	0.7979	0.8008