Project 3 – Analysis **Duc Ngo – CS3130-001**

Binary Search Tree

Introduction

A binary search tree is a data structure that quickly allows us to maintain a sorted list of numbers and can provide for efficient search. It is a node-base tree data structure which has the following properties where: the left subtree of a node contains only nodes with keys lesser than the node's key, and the right subtree of a node contains only nodes with keys greater than the node's key.

Definition

Basic Operations:

- **Search** - Find a node the tree with value n.
- **Insert** - Add a node the tree with value n.
- **Delete** - Delete a node the tree with value n.
- **Successor** - Replace the deleted node with the successor node (the smaller node in the right subtree of the node to be deleted).
- **Height** – Get the height of the binary search tree.
- **Pre-order Traversal** - Traverses a tree in a pre-order manner.
- **In-order Traversal** - Traverses a tree in an in-order manner.
- **Post-order Traversal** - Traverses a tree in a post-order manner.

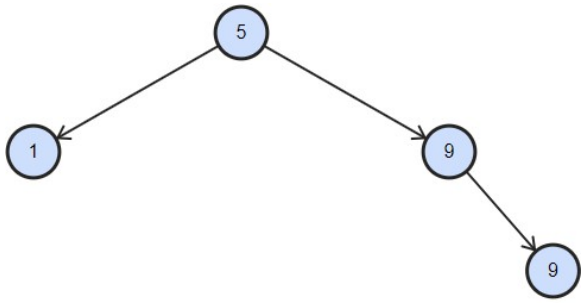
Objective

Builds *t* binary search tree by inserting *N* random keys into an initially empty tree, and then finds the tree height for N=100, 500 and 1000; and t=5, 10, 15. Find the average height of binary search trees for each pair of values of *t* and *N*.

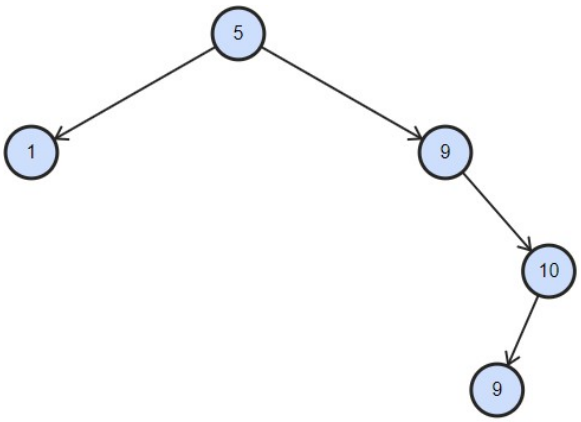
Notice

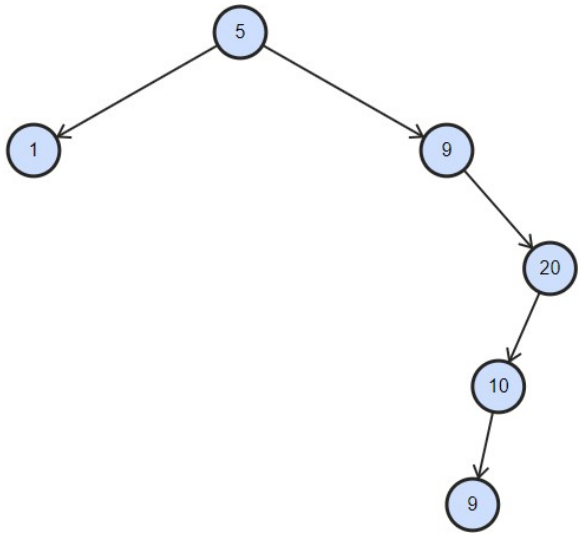
Duplicate elements are kept (must not violate the in-order traversals) and two likely scenario can happens:

- When two or more duplicates are inserted at next each other, they will always be inserted to the right side of the node. For example: inserting 9 and 9.



- When two or more duplicates are inserted and if the numbers in between the duplicate are greater than the duplicate number, they will always be inserted to the left side of the node. For example: inserting 9, 10, and 9 or inserting 9, 10, 20, 9.





Basic Computer Information (that was use to run BST height algorithms)

CPU: 2.5 GHz Quad-core Intel Core i5-7300HQ
GPU: NVDIA GeForce GTX 1050 Ti
RAM: 16GB DDR4
Storage: 128GB SSD and 1TB HDD

Experimental Data

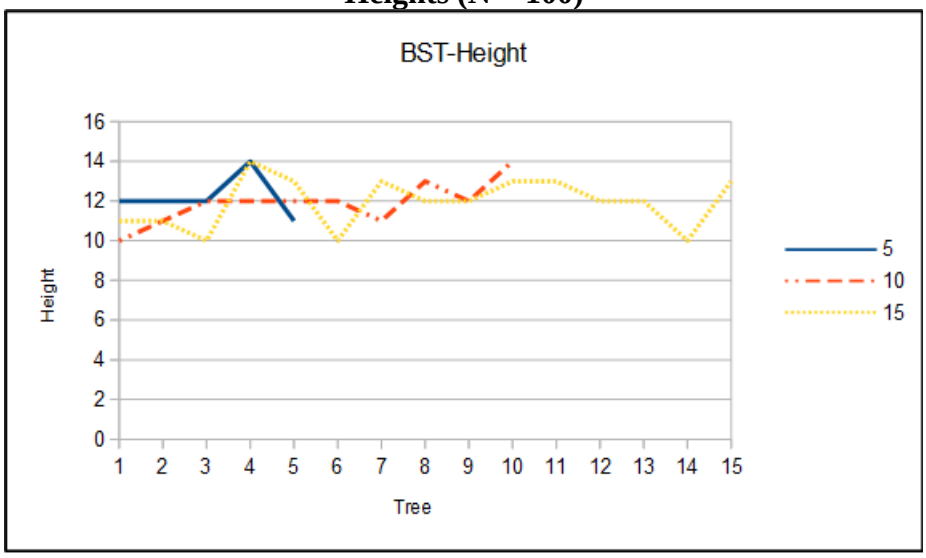
Tree Height (N = 100)

100			
Tree	**5**	**10**	**15**
1	12	10	11
2	12	11	11
3	12	12	10
4	14	12	14
5	11	12	13
6		12	10
7		11	13
8		13	12
9		12	12
10		14	13
11			13
12			12
13			12
14			10
15			13
Average	12.2	11.9	11.93333333

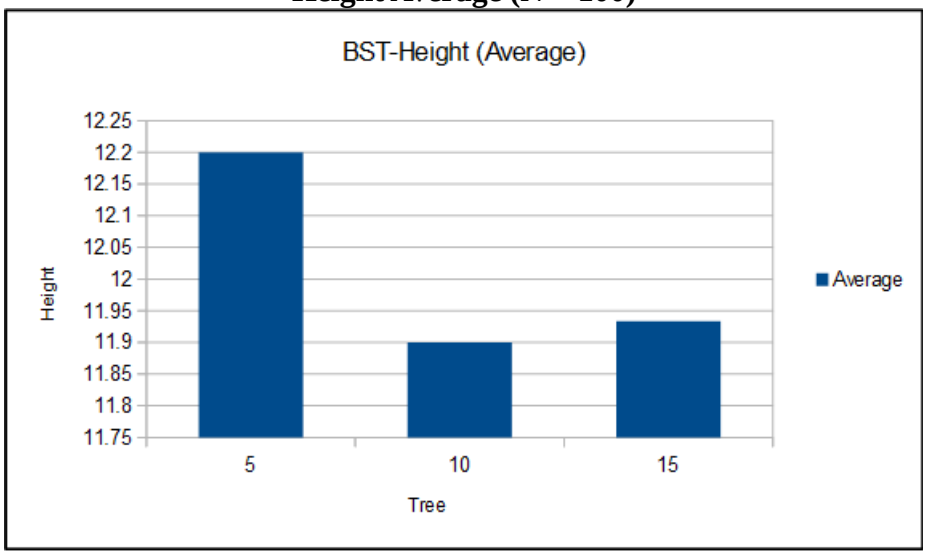
Time (N = 100)

100			
Tree	**5**	**10**	**15**
1	9.01E-004	9.42E-004	9.76E-004
2	7.92E-005	6.32E-004	1.67E-004
3	7.55E-005	1.75E-004	1.44E-004
4	1.83E-004	1.11E-004	6.93E-005
5	5.13E-005	4.92E-005	8.00E-005
6		3.98E-005	7.14E-005
7		2.13E-005	3.20E-005
8		2.63E-005	2.79E-005
9		2.17E-005	2.09E-005
10		2.09E-005	2.17E-005
11			2.22E-005
12			1.76E-005
13			2.13E-005
14			2.54E-005
15			2.42E-005
Average	2.58E-004	3.82E-004	2.87E-004

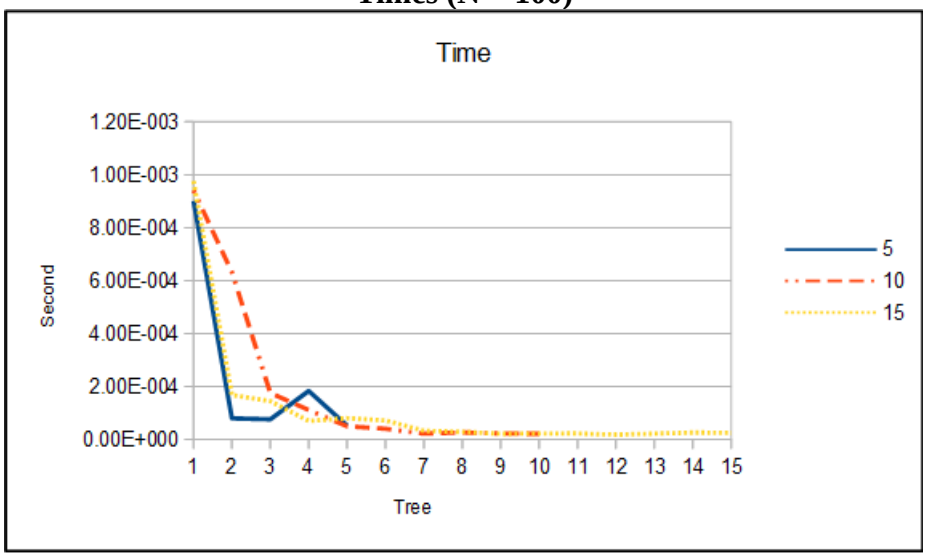
Heights (N = 100)



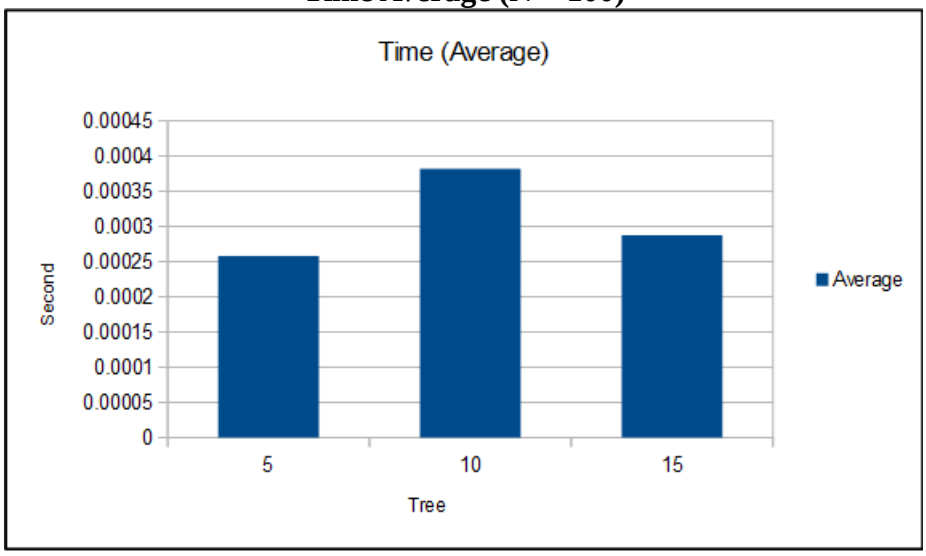
Height Average (N = 100)



Times (N = 100)



Time Average (N = 100)



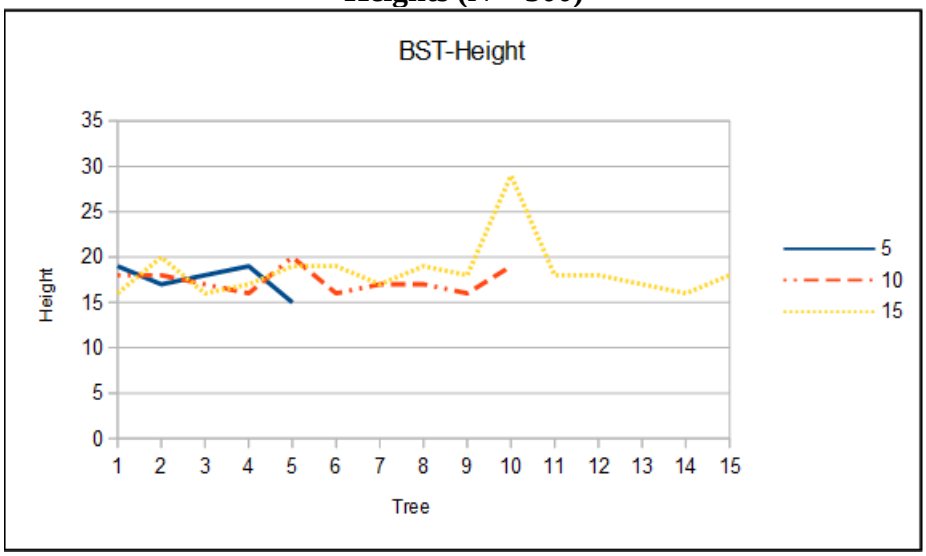
Tree Height (N = 500)

500			
Tree	**5**	**10**	**15**
1	19	18	16
2	17	18	20
3	18	17	16
4	19	16	17
5	15	20	19
6		16	19
7		17	17
8		17	19
9		16	18
10		19	29
11			18
12			18
13			17
14			16
15			18
Average	17.6	17.4	18.46666667

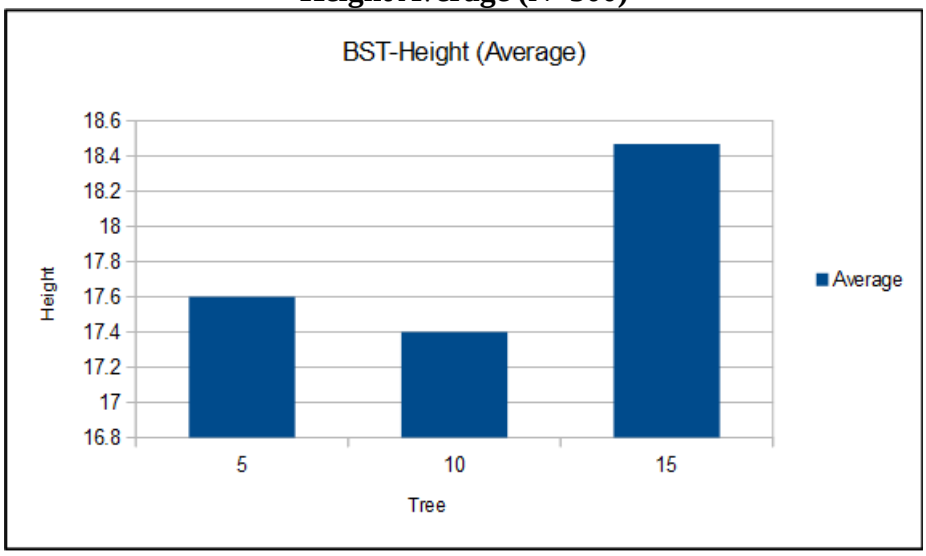
Time (N = 500)

500			
Tree	**5**	**10**	**15**
1	0.001294769	0.001215179	0.001289026
2	1.14E-004	1.17E-004	1.15E-004
3	1.01E-004	9.93E-005	9.52E-005
4	1.06E-004	1.01E-004	9.64E-005
5	1.54E-004	1.33E-004	1.36E-004
6		1.00E-004	9.68E-005
7		1.01E-004	9.76E-005
8		9.89E-005	9.52E-005
9		9.64E-005	1.01E-004
10		9.19E-005	9.31E-005
11			9.64E-005
12			1.47E-004
13			8.86E-005
14			1.12E-004
15			7.63E-005
Average	0.000354052	0.000215385	0.000182455

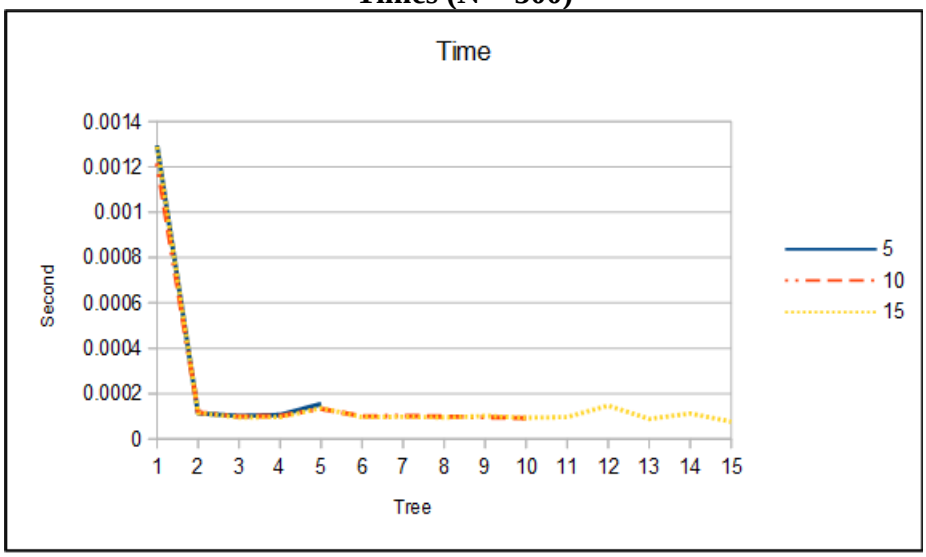
Heights (N = 500)



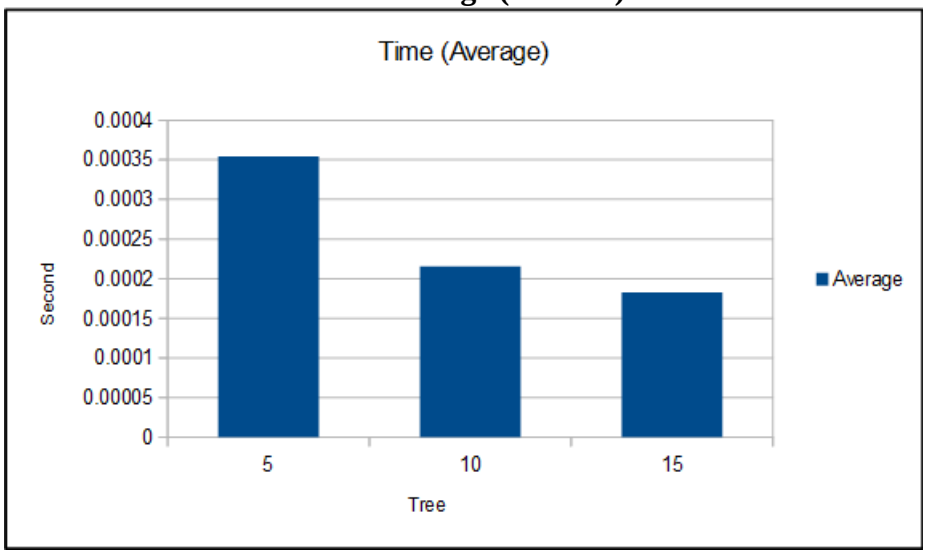
Height Average (N=500)



Times (N = 500)



Time Average (N = 500)



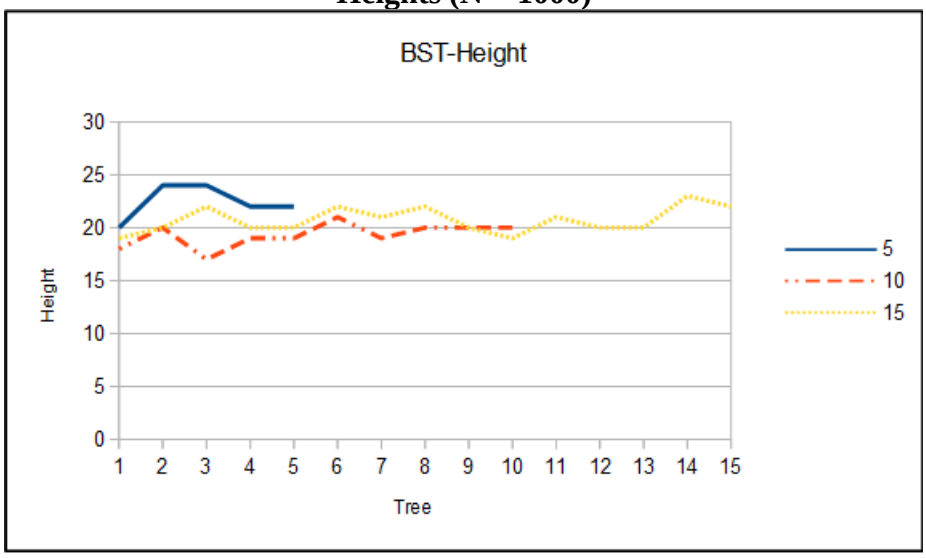
Tree Height (N = 1000)

1000			
Tree	**5**	**10**	**15**
1	20	18	19
2	24	20	20
3	24	17	22
4	22	19	20
5	22	19	20
6		21	22
7		19	21
8		20	22
9		20	20
10		20	19
11			21
12			20
13			20
14			23
15			22
Average	22.4	19.3	20.73333333

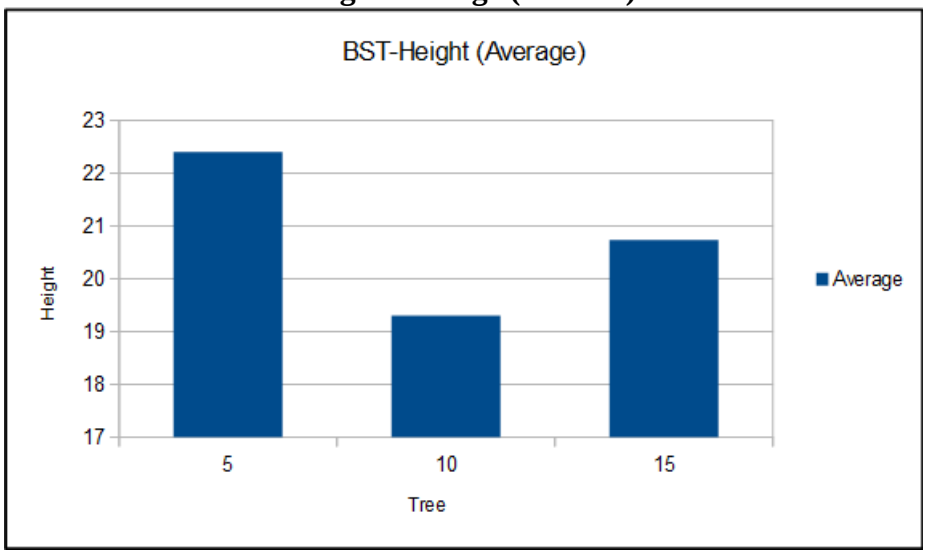
Time (N = 1000)

1000			
Tree	**5**	**10**	**15**
1	0.001682051	0.001561025	0.001827282
2	1.74E-004	2.48E-004	2.36E-004
3	2.19E-004	2.00E-004	1.96E-004
4	1.53E-004	1.91E-004	1.90E-004
5	1.76E-004	1.71E-004	1.90E-004
6		2.86E-004	2.02E-004
7		2.06E-004	2.01E-004
8		2.01E-004	1.92E-004
9		1.72E-004	1.93E-004
10		1.82E-004	1.90E-004
11			1.93E-004
12			1.99E-004
13			1.92E-004
14			2.38E-004
15			1.76E-004
Average	0.00048082	0.000341785	0.00030761

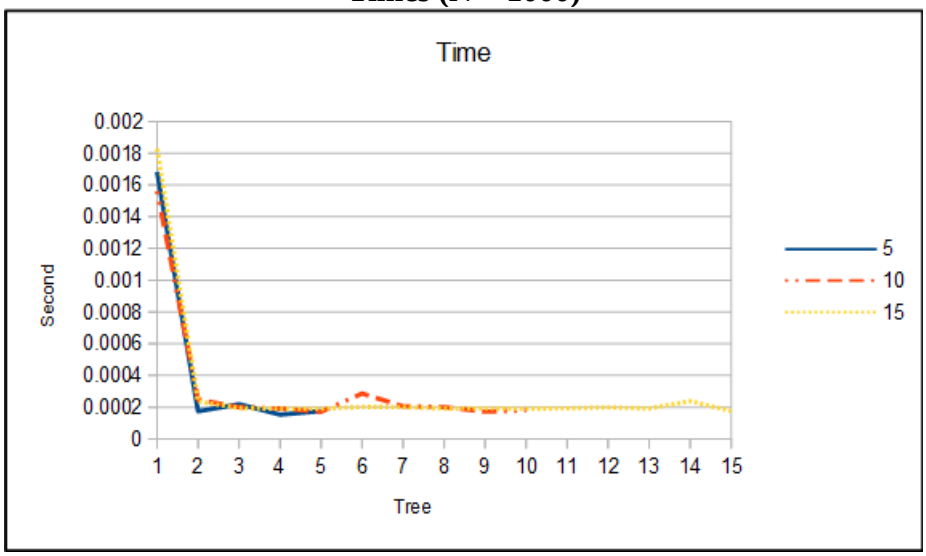
Heights (N = 1000)



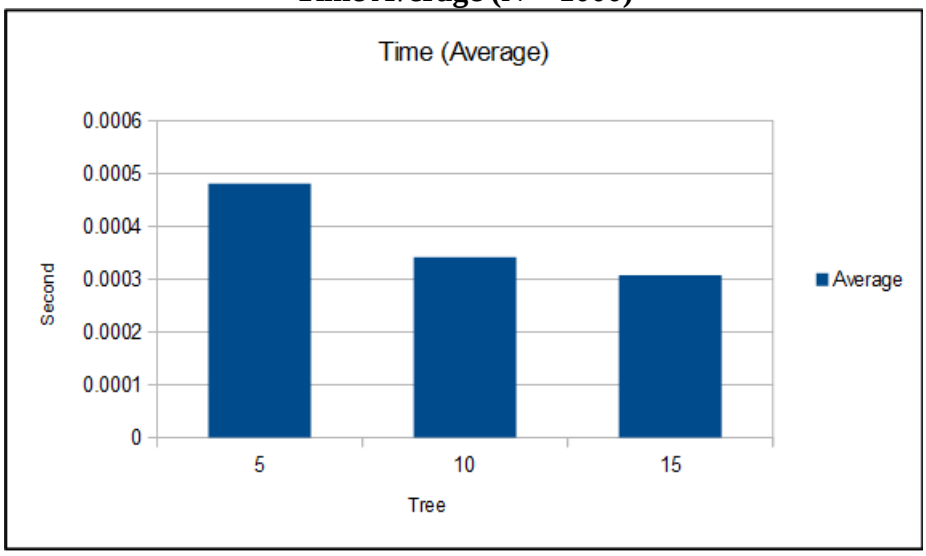
Height Average (N=1000)



Times (N = 1000)



Time Average (N = 1000)



Conclusion

Base on the average time graph (above):

- Nine average height of BST (Note: some are round up and down):
  - (100, 5) : 12
  - (100, 10) : 12
  - (100, 15) : 12
  - (500, 5) : 18
  - (500, 10) : 18
  - (500, 15) : 18
  - (1000, 5) : 22
  - (1000, 10) : 20
  - (1000, 15) : 21
- The time complexities of getting height of BST is O(n)