Hadoop Technology:

Apache Hadoop is an open-source software framework used for distributed storage
and processing of dataset of big data using the MapReduce programming model.

Hadoop is an open source implementation of Google MapReduce , GFS (Distributed File
System).Hadoop was created by Doug Cutting, the creator of Apache Lucene, the widely used
text search library.2002-2004 Doug Cutting and Mike Cafarella started working on Nutch.2003-
2004 Google publishes GFS and MapReduce paper.2004 Doug Cutting adds DFS and
MapReduce support to Nutch.

Hadoop designed and built on two independent frameworks

HDFS- Hadoop Distributed File System is reliable distributed file system that
provides high throughput access to data.

Map Reduce- Map Reduce is framework for performing high performance
distributed data processing using the divide and aggregate programming paradigm.

System Architecture
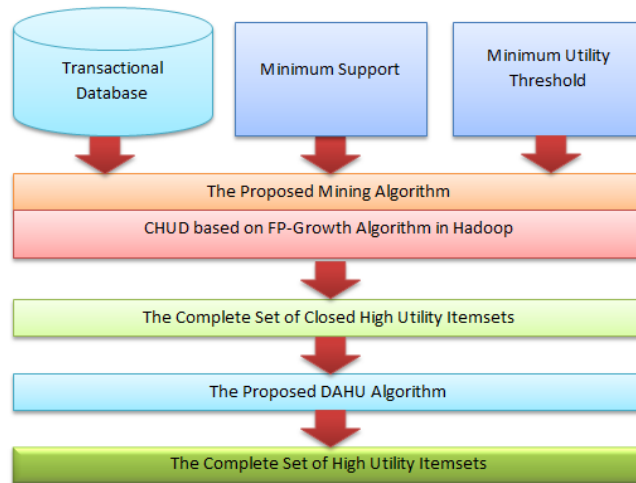
In above Figure shows the inputs Transaction Database, Minimum Support and Minimum Utility Threshold are given to the CHUD algorithm, which using FP-Growth algorithm in hadoop. Then after that algorithm, which generate the Complete Set Of High Utility Itemsets.Futher a method named DAHU (Derive All High Utility Itemsets) is proposed to generate the Complete Set High Utility itemsets.

Result Analysis:

The system was developed using Java platform and Hadoop technology. The high utility itemsets mining an algorithm called CHUD and DAHU method based on the concept of closed pattern. CHUD conduct mining Closed+ High Utility Itemsets based on pattern generated from FP algorithm in Hadoop.

Testing was performed on the Mashroom datasets with different User Utility Threshold and result obtained overall better than existing CHUD algorithm. The graph of different user utility threshold with existing CHUD algorithm vs. improve CHUD algorithm which is our implemented algorithm in hadoop is shown fig.
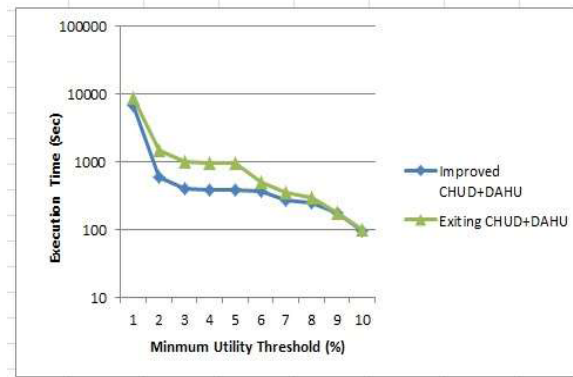
Input And Output:

When you run the project GUI will appear. In GUI you have to first browse the Dataset file, enter the Minimum Support and Minimum Utility.
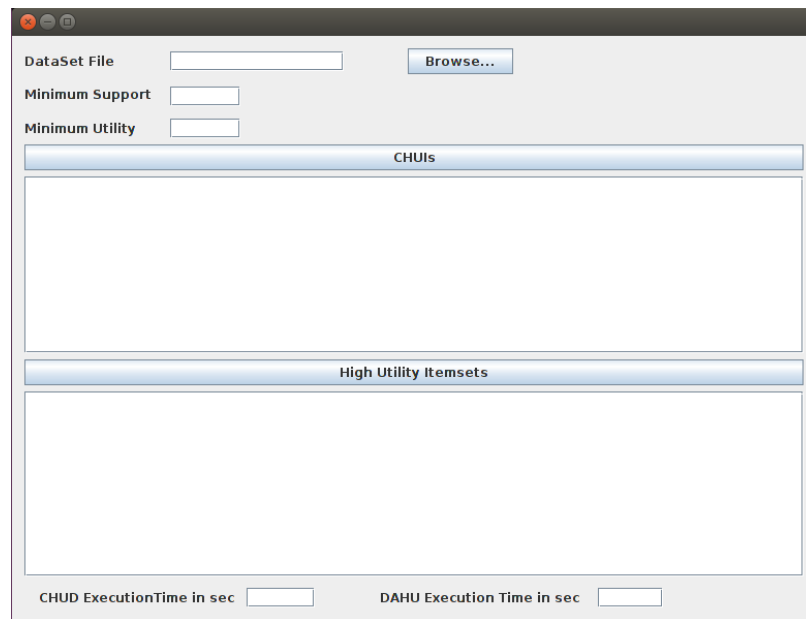
Output Screen

After entry of all details you press the CHUIs button. After that press CHUIs button. CHUIs sets will appear in upper half pat of screen from Partition0 to Partition8.Screen show left hand side are sets of item sets and right hand side it will shows the transaction utility of item sets. It also show the execution time in second at the bottom of the window(Sets of CHUIs)

Minimum Support   50

Minimum Utility   100

CHUIs

Partition :0

2,6,11,16,23,28,34,36,39,41,53,56,59,63,67,78,85,86,90,93,99,110,116	556.0
2,6,11,13,23,28,34,36,39,41,53,56,59,63,69,76,85,86,90,93,98,110,116	537.0
2,3,11,13,23,28,34,36,39,41,53,56,59,63,68,77,85,86,90,93,98,110,116	531.0
2,3,11,13,23,28,34,36,39,41,53,56,59,63,68,76,85,86,90,93,98,110,116	519.0
2,6,11,16,24,28,34,37,39,45,53,54,59,63,67,76,85,86,90,94,98,107,114	516.0
2,3,10,16,23,28,34,36,39,41,53,56,59,63,68,76,85,86,90,93,99,111,116	509.0
2,6,9,14,23,27,34,37,38,41,53,56,59,63,67,76,85,86,90,93,100,110,116	506.0
2,6,9,14,23,27,34,37,38,43,53,56,59,63,67,76,85,86,90,93,100,110,116	503.0
2,6,11,16,24,28,34,37,39,45,53,54,60,63,67,76,85,86,90,94,98,109,114	501.0
2,3,11,13,23,28,34,36,39,41,53,56,59,63,68,77,85,86,90,93,99,111,116	500.0

High Utility Itemsets

CHUD ExecutionTime in sec   13.982   **DAHU Execution Time in sec**

Sets of CHUI

After displaying the set of CHUIs on upper half part on the screen then press the High Utility Item sets button. The High Utility Item sets will display in lower half screen. It will show the sets of Item sets in left hand side and transaction utility in right hand side. It also show the execution time in seconds at the bottom of the window(High Utility Item sets)

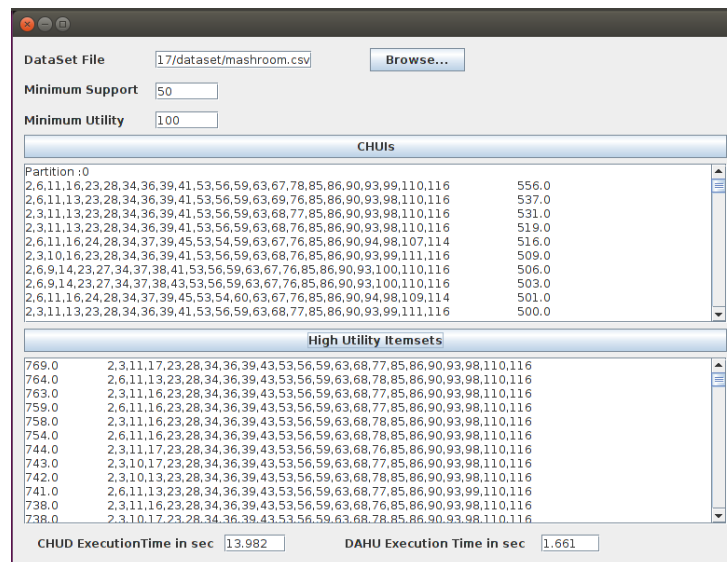**DataSet File**   17/dataset/mashroom.csv   Browse...

Minimum Support   50

Minimum Utility   100

CHUIs

Partition :0

2,6,11,16,23,28,34,36,39,41,53,56,59,63,67,78,85,86,90,93,99,110,116	556.0
2,6,11,13,23,28,34,36,39,41,53,56,59,63,69,76,85,86,90,93,98,110,116	537.0
2,3,11,13,23,28,34,36,39,41,53,56,59,63,68,77,85,86,90,93,98,110,116	531.0
2,3,11,13,23,28,34,36,39,41,53,56,59,63,68,76,85,86,90,93,98,110,116	519.0
2,6,11,16,24,28,34,37,39,45,53,54,59,63,67,76,85,86,90,94,98,107,114	516.0
2,3,10,16,23,28,34,36,39,41,53,56,59,63,68,76,85,86,90,93,99,111,116	509.0
2,6,9,14,23,27,34,37,38,41,53,56,59,63,67,76,85,86,90,93,100,110,116	506.0
2,6,9,14,23,27,34,37,38,43,53,56,59,63,67,76,85,86,90,93,100,110,116	503.0
2,6,11,16,24,28,34,37,39,45,53,54,60,63,67,76,85,86,90,94,98,109,114	501.0
2,3,11,13,23,28,34,36,39,41,53,56,59,63,68,77,85,86,90,93,99,111,116	500.0

High Utility Itemsets

769.0	2,3,11,17,23,28,34,36,39,43,53,56,59,63,68,77,85,86,90,93,98,110,116
764.0	2,6,11,13,23,28,34,36,39,43,53,56,59,63,68,78,85,86,90,93,98,110,116
763.0	2,3,11,16,23,28,34,36,39,43,53,56,59,63,68,77,85,86,90,93,98,110,116
759.0	2,6,11,16,23,28,34,36,39,43,53,56,59,63,68,77,85,86,90,93,98,110,116
758.0	2,3,11,16,23,28,34,36,39,43,53,56,59,63,68,78,85,86,90,93,98,110,116
754.0	2,6,11,16,23,28,34,36,39,43,53,56,59,63,68,78,85,86,90,93,98,110,116
744.0	2,3,11,17,23,28,34,36,39,43,53,56,59,63,68,76,85,86,90,93,98,110,116
743.0	2,3,10,17,23,28,34,36,39,43,53,56,59,63,68,77,85,86,90,93,98,110,116
742.0	2,3,10,13,23,28,34,36,39,43,53,56,59,63,68,78,85,86,90,93,98,110,116
741.0	2,6,11,13,23,28,34,36,39,43,53,56,59,63,68,77,85,86,90,93,99,110,116
738.0	2,3,11,16,23,28,34,36,39,43,53,56,59,63,68,76,85,86,90,93,98,110,116
738.0	2,3,10,17,23,28,34,36,39,43,53,56,59,63,68,78,85,86,90,93,98,110,116

CHUD ExecutionTime in sec   13.982   **DAHU Execution Time in sec**   1.661

High Utility Item sets