A Study on Comparison of Bayesian Network Structure Learning Algorithm for Selecting Appropriate Model

Jae-seong Yoo

Dept. of Statistics

December 5, 2014

Title

Goal

- In this paper, we compare the performance between the Bayesian network structure learning algorithms provided by **bnlearn** package in **R**.

- The performance is evaluated by
    - using a score

      or
    - comparing between the target network and the learnt network.

  In this paper, it was confirmed that algorithm specific performance test results using fore-mentioned methods are different.

- A data generator based on Bayesian network model using **R** is built and introduced.

- The aim of this paper is to provide objective guidance of selecting suitable algorithm in accordance to target network using synthetic data generated based on topology.

Bayesian Network

A BN defines a unique joint probability distribution over X given by

$$P_B(X_1, \cdots, X_n) = \prod_{i=1}^{n} P_B(X_i | \prod_{X_i}).$$

- A BN encodes the independence assumptions over the component random variables of X.
- An edge (j, i) in E represents a direct dependency of X_i from X_j.
- The set of all Bayesian networks with n variables is denotes by B_n.

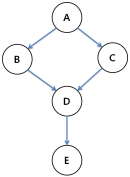


Figure: $P(A, B, C, D, E) = P(A)P(B|A)P(C|A)P(D|B, C)P(E|D)$

Bayesian Network Structure Learning

Learning a Bayesian network is as follows:

Given a data $T = \{y_1, \cdots, y_n\}$ and a scoring function ϕ, the problem of learning a Bayesian network is to find a Bayesian network $B \in B_n$ that maximizes the value $\phi(B, T)$.

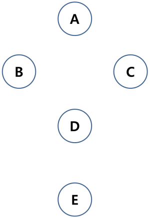


Figure: A model before learning structure

Available Constraint-based learning algorithms

Grow-Shrink (GS) based on the Grow-Shrink Markov Blanket, the first (and simplest) Markov blanket detection algorithm used in a structure learning algorithm.

Incremental Association (IAMB) based on the Markov blanket detection algorithm of the same name, which is based on a two-phase selection scheme (a forward selection followed by an attempt to remove false positives).

Available Score-based Learning Algorithms

Hill-Climbing (HC) a hill climbing greedy search on the space of the directed graphs. The optimized implementation uses score caching, score decomposability and score equivalence to reduce the number of duplicated tests.

Tabu Search (TABU) a modified hill climbing able to escape local optima by selecting a network that minimally decreases the score function.

Available Hybrid Learning Algorithms

Max-Min Hill-Climbing (MHHC) a hybrid algorithm which combines the Max-Min Parents and Children algorithm (to restrict the search space) and the Hill-Climbing algorithm (to find the optimal network structure in the restricted space).

Restricted Maximization (RSMAX2) a more general implementation of the Max-Min Hill-Climbing, which can use any combination of constraint-based and score-based algorithms.

The Number of Graphical Errors in the Learnt Structure

In terms of the number of graphical errors in the learnt structure.

		Target Network	Learnt Network	Direction
C	(Correct Arcs)	exist	exist	correct
M	(Missing Arcs)	exist	not exist	
WO	(Wrongly Oriented Arcs)	exist	exist	wrong
WC	(Wrongly Corrected Arcs)	not exist	exist	

Network Scores

In all four cases, the higher the value of the metric, the better the network.

BDe $BDe(B,T) = P(B,T) = P(B) \times \prod_{i=1}^{n} \prod_{j=1}^{q_i} \left(\frac{\Gamma(N'_{ij})}{\Gamma(N_{ij}+N'_{ij})} \times \prod_{k=1}^{r_i} \frac{\Gamma(N_{ijk}+N'_{ijk})}{\Gamma(N'_{ijk})} \right)$

$$\phi(B|T) = LL(B|T) - f(N)|B|,$$

Log-Likelihood(LL) If $f(N) = 0$, we have the **LL** score.

AIC If f(N) = 1, we have the **AIC** scoring function:

BIC If $f(N) = \frac{1}{2}\log(N)$, we have the **BIC** score.

Data Generation with BN_Data_Generator in R

BN_Data_Generator {User-Defined Function}

Description It based on a Bayesian network model to generates synthetic data.

Usage BN_Data_Generator (arcs, input_Probs, n, node_names)

URL https://github.com/praster1/BN_Data_Generator

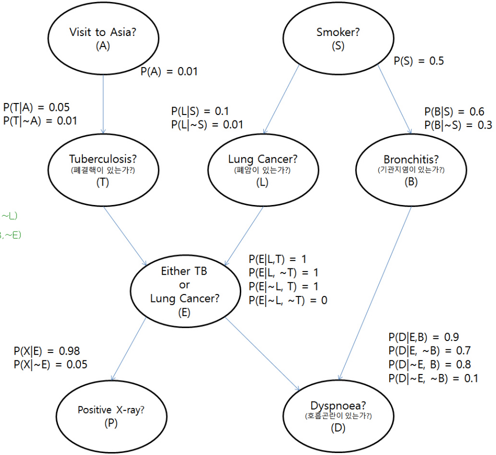
Arguments

arcs	(matrix)	A matrix that determines the arcs.
input_Probs	(list)	The conditional probabilities.
n	(constant)	Sample Size
node_names	(vector)	Node names

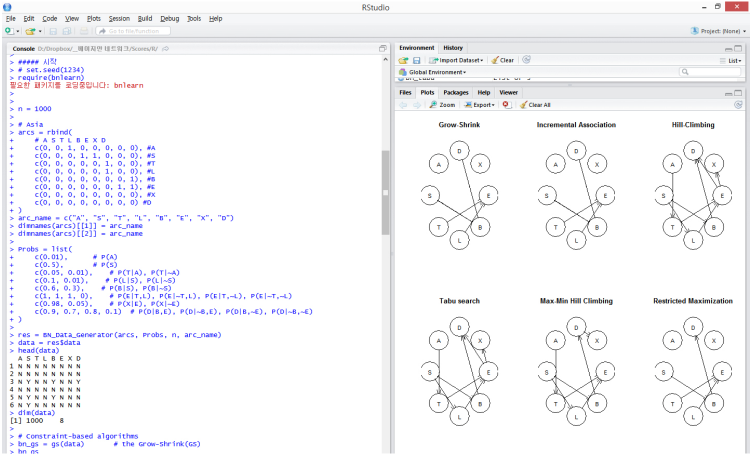
Data Generation with BN_Data_Generator in R

```r
# Asia
arcs = rbind(
      #  A  S  T  L  B  E  X  D
      c(0,  0, 1, 0, 0, 0, 0, 0), #A
      c(0,  0, 0, 1, 1, 0, 0, 0), #S
      c(0,  0, 0, 0, 0, 1, 0, 0), #T
      c(0,  0, 0, 0, 0, 1, 0, 0), #L
      c(0,  0, 0, 0, 0, 0, 0, 1), #B
      c(0,  0, 0, 0, 0, 0, 1, 1), #E
      c(0,  0, 0, 0, 0, 0, 0, 0), #X
      c(0,  0, 0, 0, 0, 0, 0, 0)  #D
      )
arc_name = c("A", "S", "T", "L", "B", "E", "X", "D")
dimnames(arcs)[[1]] = arc_name
dimnames(arcs)[[2]] = arc_name

Probs = list(
    c(0.01),                    # P(A)
    c(0.5),                     # P(S)
    c(0.05, 0.01),              # P(T|A), P(T|~A)
    c(0.1, 0.01),               # P(L|S), P(L|~S)
    c(0.6, 0.3),              # P(B|S), P(B|~S)
    c(1, 1, 1, 0),              # P(E|T,L), P(E|~T,L), P(E|T,~L), P(E|~T,~L)
    c(0.98, 0.05),              # P(X|E), P(X|~E)
    c(0.9, 0.7, 0.8, 0.1)       # P(D|B,E), P(D|~B,E), P(D|B,~E), P(D|~B,~E)
)
```



Data Generation with BN_Data_Generator in R



Outline

Prerequisite

- All experiments are repeated 100 times, and overall results are reported.
- Constraint-based Learning Algorithms often makes undirected arcs. So, this has been excluded from comparison.

Asia DataSet

Description Small synthetic data set from Lauritzen and Spiegelhalter (1988) about lung diseases (tuberculosis, lung cancer or bronchitis) and visits to Asia.

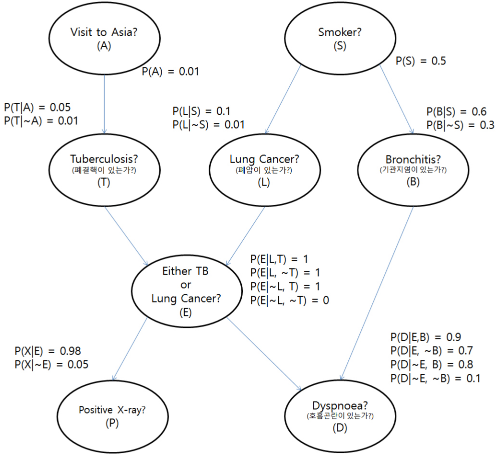
Number of nodes 8

Number of arcs 8

Number of parameters 18

Source Lauritzen S, Spiegelhalter D (1988).
"Local Computation with Probabilities on Graphical Structures and their Application to Expert Systems (with discussion)".
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 50(2), 157-224.

Asia DataSet



Insurance DataSet

Description Insurance is a network for evaluating car insurance risks.

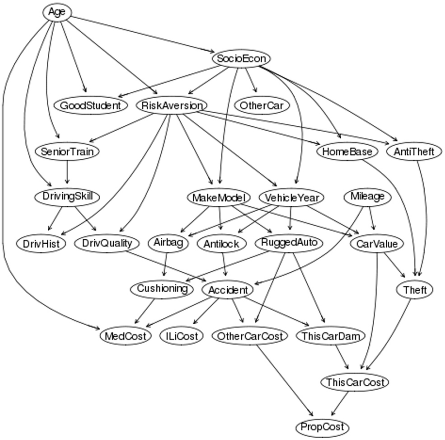
Number of nodes 27

Number of arcs 52

Number of parameters 984

Source Binder J, Koller D, Russell S, Kanazawa K (1997). "Adaptive Probabilistic Networks with Hidden Variables". Machine Learning, 29(2-3), 213-244.

Insurance DataSet



Alarm DataSet

Description The ALARM ("A Logical Alarm Reduction Mechanism") is a Bayesian network designed to provide an alarm message system for patient monitoring.

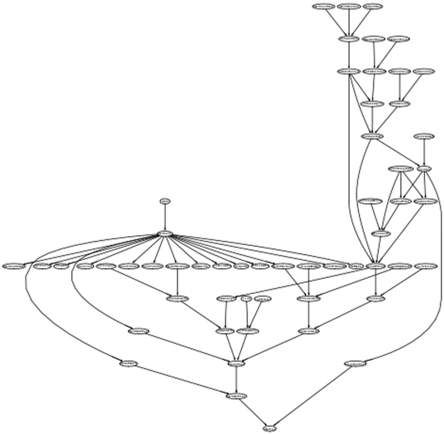
Number of nodes 37

Number of arcs 46

Number of parameters 509

Source Beinlich I, Suermondt HJ, Chavez RM, Cooper GF (1989).
"The ALARM Monitoring System: A Case Study with Two Probabilistic Inference Techniques for Belief Networks."
In "Proceedings of the 2nd European Conference on Artificial Intelligence in Medicine", pp. 247-256. Springer-Verlag.

Alarm DataSet



HailFinder DataSet

Description Hailfinder is a Bayesian network designed to forecast severe summer hail in northeastern Colorado.

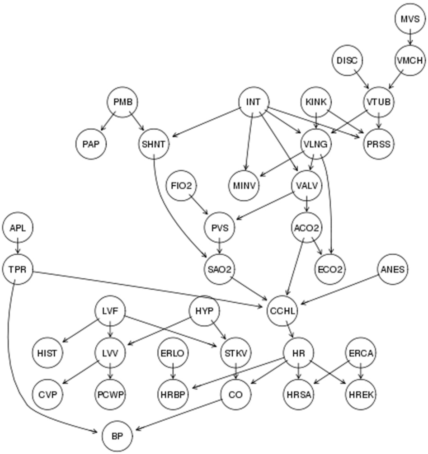
Number of nodes 56

Number of arcs 66

Number of parameters 2656

Source Abramson B, Brown J, Edwards W, Murphy A, Winkler RL (1996). "Hailfinder: A Bayesian system for forecasting severe weather". International Journal of Forecasting, 12(1), 57-71.

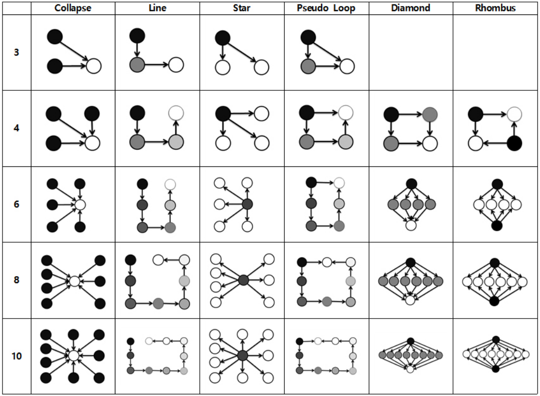
HailFinder DataSet



Summary

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
Asia	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	3	2	1	4
Insurance	2	1	3	4	2	1	3	4	3	4	2	1	2	1	4	3	1	2	4	3
Alarm	2	1	3	4	2	1	3	4	3	4	2	1	1	2	3	4	1	2	3	4
HaliFinder	2	1	3	4	2	1	3	4	4	4	2	1	1	2	3	4	1	2	3	4
Sample Size 5000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
Asia	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	3	2	1	4
Insurance	2	1	3	4	2	1	3	4	3	4	2	1	1	2	3	4	1	2	3	4
Alarm	1	2	3	4	2	1	3	4	4	3	2	1	1	3	2	4	2	3	1	4
HaliFinder	1	1	3	4	1	1	3	4	4	4	2	1	4	4	4	4	2	2	1	4
Sample Size 10000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
Asia	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	3	1	2	4
Insurance	2	1	3	4	2	1	3	4	3	4	2	1	1	3	2	4	1	2	3	4
Alarm	2	1	3	4	2	1	3	4	4	4	2	1	1	2	3	4	1	2	3	4
HaliFinder	2	1	3	4	1	2	3	4	4	3	2	1	2	1	4	4	3	2	1	4

Varying topologies and number of nodes

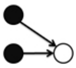
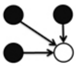


Eitel J. M. Lauría,
"An Information-Geometric Approach to Learning Bayesian Network Topologies from Data",
Innovations in Bayesian Networks Studies in Computational Intelligence Volume 156, 2008, pp 187-217

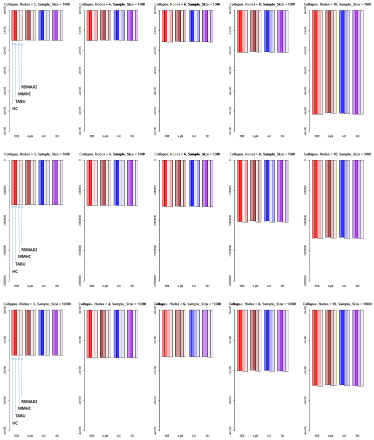
Prerequisite

- Cardinality was limited to two.
- The probability value, which is imparted optionally under U(0, 1) distribution.
- All experiments are repeated 100 times, and overall results are reported.
- Constraint-based Learning Algorithms often makes undirected arcs. So, this has been excluded from comparison.

Collapse

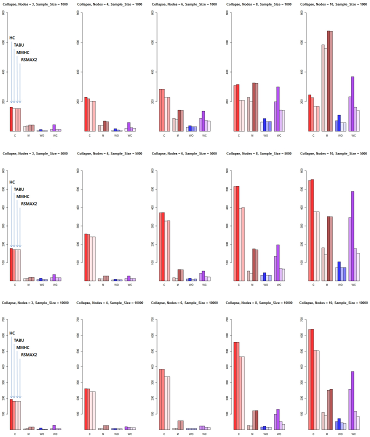
	3	4	6	8	10
Collapse					

Collapse (Score)



*The higher the value, the color is dark.

Collapse (Arcs)

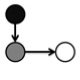
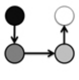
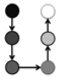
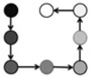
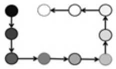


*The higher the value, the color is dark.

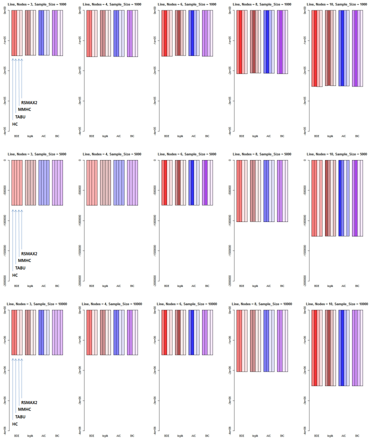
Collapse

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	2	1	4	4	1	4	2	2	4	3	1	1	4	1	4	4	4	1	4	4
4	2	1	4	3	1	2	4	3	3	4	1	2	4	1	2	4	4	1	2	4
6	2	1	4	3	1	1	4	3	3	4	1	2	4	1	2	3	2	1	3	4
8	2	1	3	4	2	1	4	4	3	4	1	2	4	1	3	2	2	1	3	4
10	2	1	3	4	1	2	4	3	3	4	1	2	2	1	4	4	2	1	3	4
Sample Size 5000	**Score**				**C**				**M**				**WO**				**WC**			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	2	1	4	4	1	2	4	4	4	3	1	1	4	1	4	4	4	1	4	4
4	2	1	4	4	1	2	4	4	3	4	1	1	4	1	2	2	4	1	4	4
6	2	1	4	3	2	1	4	3	3	4	1	2	2	1	4	2	2	1	3	4
8	2	1	4	3	2	1	4	3	3	4	1	2	2	1	4	2	2	1	3	4
10	2	1	3	4	2	1	4	4	3	4	1	2	4	1	3	2	2	1	3	4
Sample Size 10000	**Score**				**C**				**M**				**WO**				**WC**			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	2	1	4	4	1	2	4	4	4	3	1	1	4	1	4	4	4	1	4	4
4	2	1	4	4	1	2	4	4	4	4	1	1	2	1	4	4	1	2	4	4
6	1	1	3	4	1	1	4	4	4	4	1	1	1	1	1	1	1	1	3	4
8	2	1	3	4	1	2	3	4	3	4	2	1	2	1	4	4	2	1	3	4
10	2	1	3	4	2	1	3	4	3	4	2	1	2	1	3	4	2	1	3	4

Line

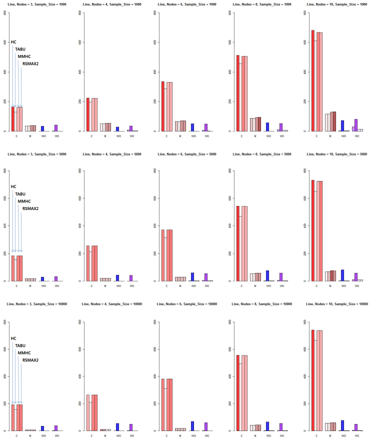
	3	4	6	8	10
Line					

Line (Score)



*The higher the value, the color is dark.

Line (Arcs)



*The higher the value, the color is dark.

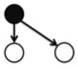
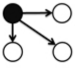
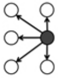
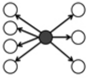
Line

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	2	1	4	4
4	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	2	1	4	4
6	2	1	3	4	1	4	2	2	3	4	1	1	4	1	4	4	2	1	3	4
8	2	1	3	4	1	4	2	3	3	4	2	1	4	1	4	4	2	1	4	4
10	2	1	3	4	1	4	2	3	4	4	2	1	4	1	4	4	2	1	4	4

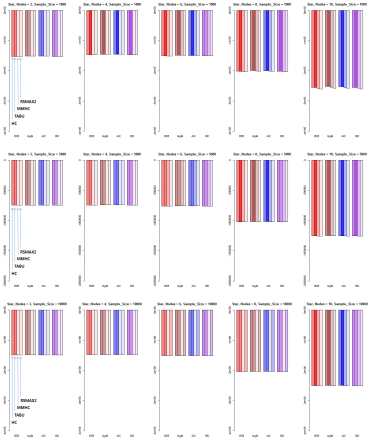
Sample Size 5000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	1	1	1	1	1	4	1	1	1	1	1	1	4	1	4	4	4	1	4	4
4	1	1	1	1	1	4	1	1	1	1	1	1	4	1	4	4	4	1	4	4
6	2	1	4	4	1	4	1	1	1	1	1	1	4	1	4	4	4	1	4	4
8	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	2	1	4	4
10	1	1	4	3	1	4	3	2	4	3	1	2	4	1	4	4	2	1	4	4

Sample Size 10000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	1	1	4	4	1	4	1	1	1	1	1	1	4	1	4	4	2	1	4	4
4	1	1	4	4	3	4	1	1	1	1	4	4	4	1	4	4	2	1	4	4
6	2	1	4	4	1	4	1	1	1	1	1	1	4	1	4	4	2	1	4	4
8	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
10	2	1	3	4	1	4	2	2	4	4	1	1	4	1	4	4	2	1	3	4

Star

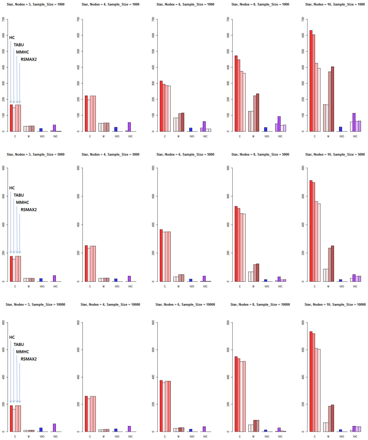
	3	4	6	8	10
Star					

Star (Score)



*The higher the value, the color is dark.

Star (Arcs)



*The higher the value, the color is dark.

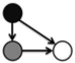
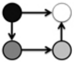
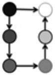
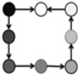
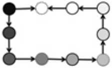
Star

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	2	1	4	4
4	2	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	2	1	4	4
6	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	2	1	4	4
8	2	1	3	4	1	2	3	4	3	4	2	1	4	1	4	4	2	1	4	3
10	2	1	3	4	1	2	3	4	3	4	2	1	4	1	4	4	4	1	3	2

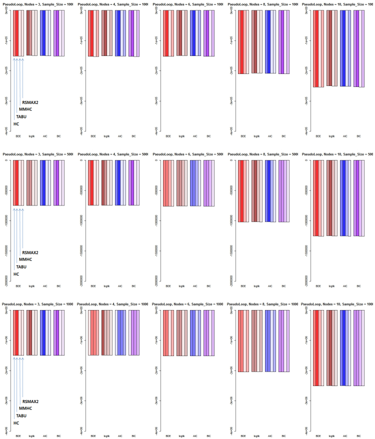
Sample Size 5000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
4	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
6	1	1	3	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
8	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	4	1	3	2
10	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	4	1	3	2

Sample Size 10000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
4	1	1	4	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
6	1	1	4	3	1	4	3	2	4	4	1	2	4	1	4	4	4	1	4	4
8	1	1	4	3	1	2	4	3	4	4	1	2	4	1	4	4	4	1	2	3
10	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	4	1	2	2

PseudoLoop

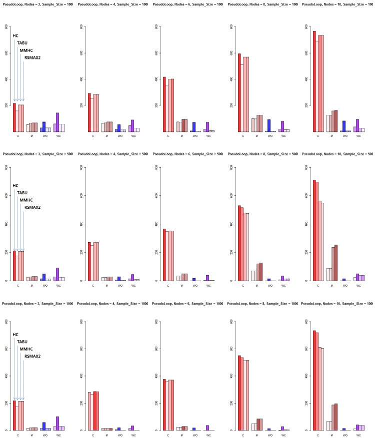
	3	4	6	8	10
Pseudo Loop					

PseudoLoop (Score)



*The higher the value, the color is dark.

PseudoLoop (Arc)



*The higher the value, the color is dark.

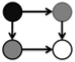
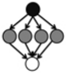
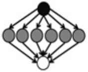
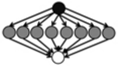
PseudoLoop

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	2	1	4	4	1	4	2	2	4	3	1	1	2	1	4	4	2	1	4	4
4	2	1	4	3	1	4	2	2	4	3	1	1	2	1	4	4	2	1	4	4
6	2	1	4	3	1	4	3	2	3	4	1	2	2	1	4	4	2	1	4	4
8	2	1	3	4	1	4	2	2	3	4	1	1	4	1	4	4	2	1	4	4
10	2	1	4	3	1	4	2	3	4	3	2	1	4	1	4	4	2	1	3	4

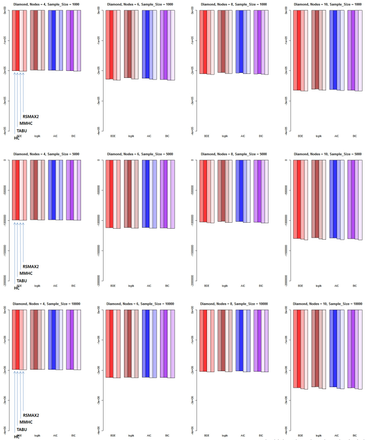
Sample Size 5000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	2	1	4	4	1	4	2	2	4	3	1	1	2	1	4	4	2	1	4	4
4	2	1	4	4	1	4	2	2	4	4	1	1	2	1	4	4	2	1	4	4
6	1	1	3	4	1	4	2	2	4	4	1	1	4	1	4	4	4	1	4	4
8	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	4	1	3	2
10	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	4	1	3	2

Sample Size 10000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
3	2	1	4	4	1	4	2	2	4	3	1	1	4	1	4	4	4	1	4	4
4	4	1	1	3	3	4	1	2	4	4	4	1	2	1	4	4	2	1	4	4
6	1	1	4	3	1	4	3	2	4	4	1	2	4	1	4	4	4	1	4	4
8	1	1	4	3	1	2	4	3	4	4	1	2	4	1	4	4	4	1	2	3
10	2	1	3	4	1	2	3	4	4	4	2	1	4	1	4	4	4	1	2	2

Diamond

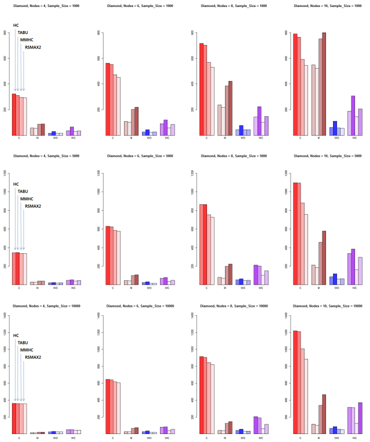
	3	4	6	8	10
Diamond					

Diamond (Score)



*The higher the value, the color is dark.

Diamond (Arc)



*The higher the value, the color is dark.

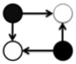
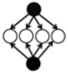
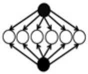
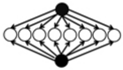
Diamond

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
4	2	1	4	3	1	2	3	4	3	4	2	1	2	1	4	4	2	1	4	3
6	2	1	3	4	1	2	3	4	3	4	2	1	2	1	4	3	2	1	4	3
8	2	1	3	4	1	2	3	4	3	4	2	1	4	1	4	4	3	1	4	2
10	2	1	3	4	1	2	3	4	3	4	2	1	2	1	3	4	3	1	4	2

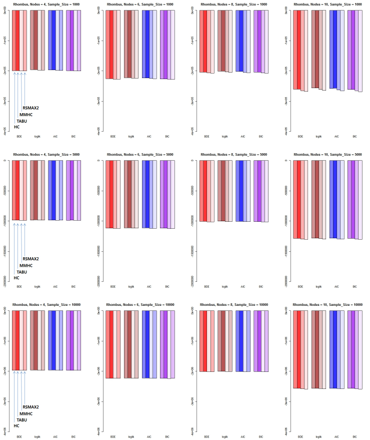
Sample Size 5000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
4	2	1	4	3	2	1	3	4	3	4	1	1	2	1	4	3	2	1	4	3
6	2	1	4	3	1	2	3	4	3	4	2	1	2	1	4	4	2	1	4	3
8	2	1	3	4	2	1	3	4	3	4	2	1	2	1	4	3	1	2	4	3
10	2	1	3	4	1	2	3	4	3	4	2	1	2	1	4	3	2	1	4	3

Sample Size 10000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
4	2	1	3	4	1	2	3	4	3	4	2	1	2	1	4	4	1	1	4	4
6	2	1	4	3	1	2	3	4	3	4	2	1	2	1	4	4	2	1	4	3
8	2	1	4	3	1	2	3	4	3	4	2	1	2	1	4	3	1	2	4	3
10	2	1	3	4	1	2	3	4	3	4	2	1	2	1	3	4	2	3	4	1

Rhombus

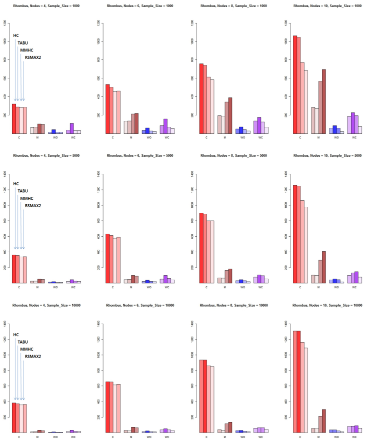
	3	4	6	8	10
Rhombus					

Rhombus (Score)



*The higher the value, the color is dark.

Rhombus (Arc)



*The higher the value, the color is dark.

Rhombus

Sample Size 1000	Score				C				M				WO				WC			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
4	2	1	4	3	1	2	4	3	4	3	1	2	4	1	4	4	2	1	4	4
6	2	1	3	4	1	2	4	3	4	3	2	1	2	1	3	4	2	1	3	4
8	2	1	3	4	1	2	3	4	3	4	2	1	2	1	3	4	2	1	3	4
10	2	1	3	4	1	2	3	4	3	4	2	1	3	1	2	4	3	1	2	4
Sample Size 5000	**Score**				**C**				**M**				**WO**				**WC**			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
4	2	1	4	3	1	2	4	3	3	4	1	2	2	1	4	4	4	1	2	4
6	2	1	3	4	1	2	4	3	4	3	1	2	3	1	2	4	3	1	2	4
8	2	1	3	4	1	2	3	4	3	4	2	1	2	1	2	4	3	1	2	4
10	2	1	3	4	1	2	3	4	3	4	2	1	3	1	2	4	3	2	1	4
Sample Size 10000	**Score**				**C**				**M**				**WO**				**WC**			
	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2	HC	TABU	MMHC	RSMAX2
4	2	1	4	3	1	2	4	3	4	3	1	2	4	1	4	4	4	1	2	4
6	2	1	4	3	1	2	4	3	3	4	1	2	2	1	2	4	2	1	2	4
8	2	1	4	3	2	1	3	4	3	4	2	1	2	1	3	4	3	1	1	4
10	2	1	3	4	2	1	3	4	3	4	2	1	1	1	3	4	3	2	1	4

Discussion

- In most cases using synthetic data according to topology,
  If comparing by score, then TABU search has good performance.
  But comparing by reference to "What C is the lot?", then HC has also good performance.

- Hybrid algorithm compared to Score-based algorithm is found to be that draw the arc more conservative.

- Sample size is larger, then C was increased.
  In addition, M, WO and WC was decreased.

- About Line and Star form, the performance difference due to relatively algorithm was not large compared to other topology.