A Generic Method For Estimating System Reliability Using Bayesian Networks

Lecturer : Amir M . Pirhosseinloo

Professor : Dr. Safabakhsh

Overview

1.Introduction

Bayesian Networks, System Reliability, Traditional Approaches, Challenges

2.K2 algorithm

Notations, Pseudo code

3.Illustrative step by step example

4.Experimental analysis

Performance, Accuracy, Error rate

5.Conclusion

Introduction

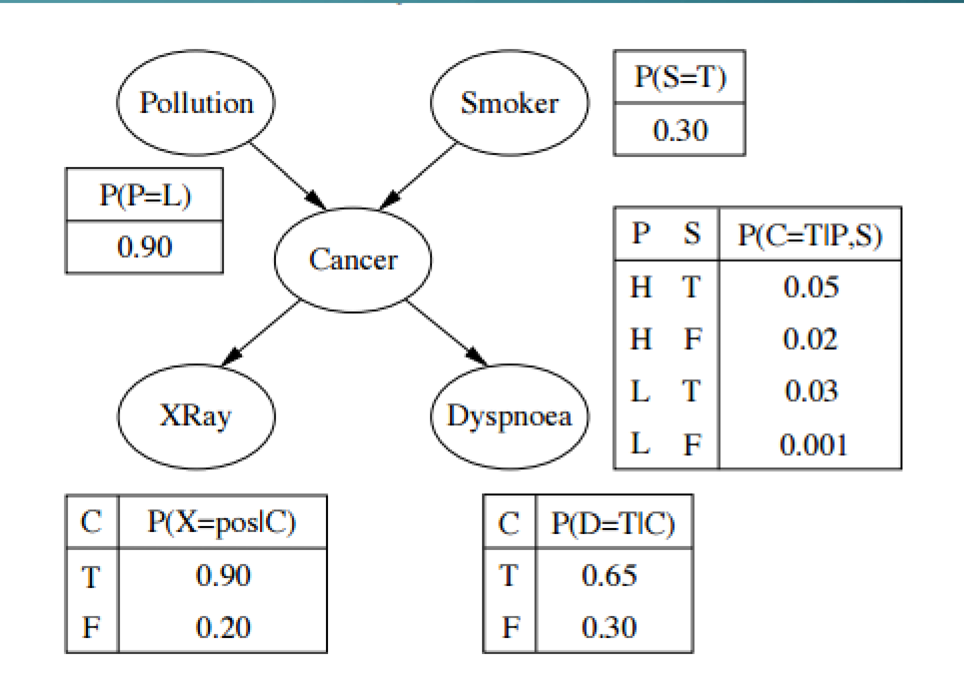
Bayesian Networks

What is a Bayesian Network ?

DAG: Directed Acyclic Graph

Nodes: System Components

Edges: Relations

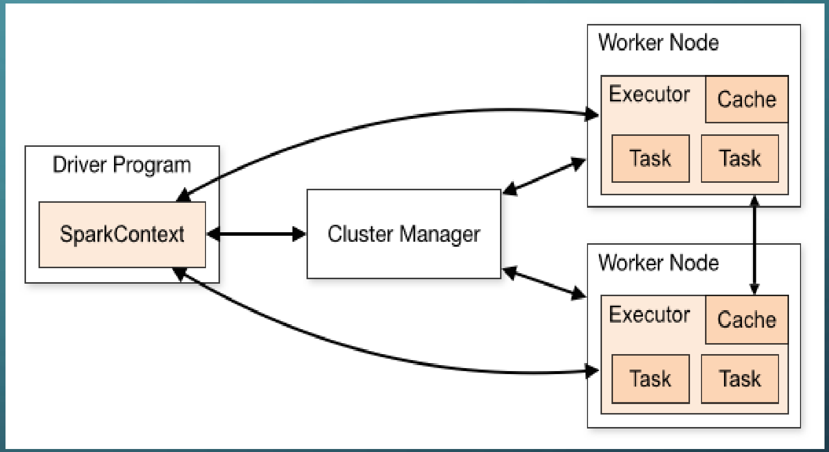


www.math.stackexchange.com/questions/1709832/confusion-in-a-simple-bayesian-network

System Reliability

- Probability
- Intended function
- Specific period of time
- Stated condition

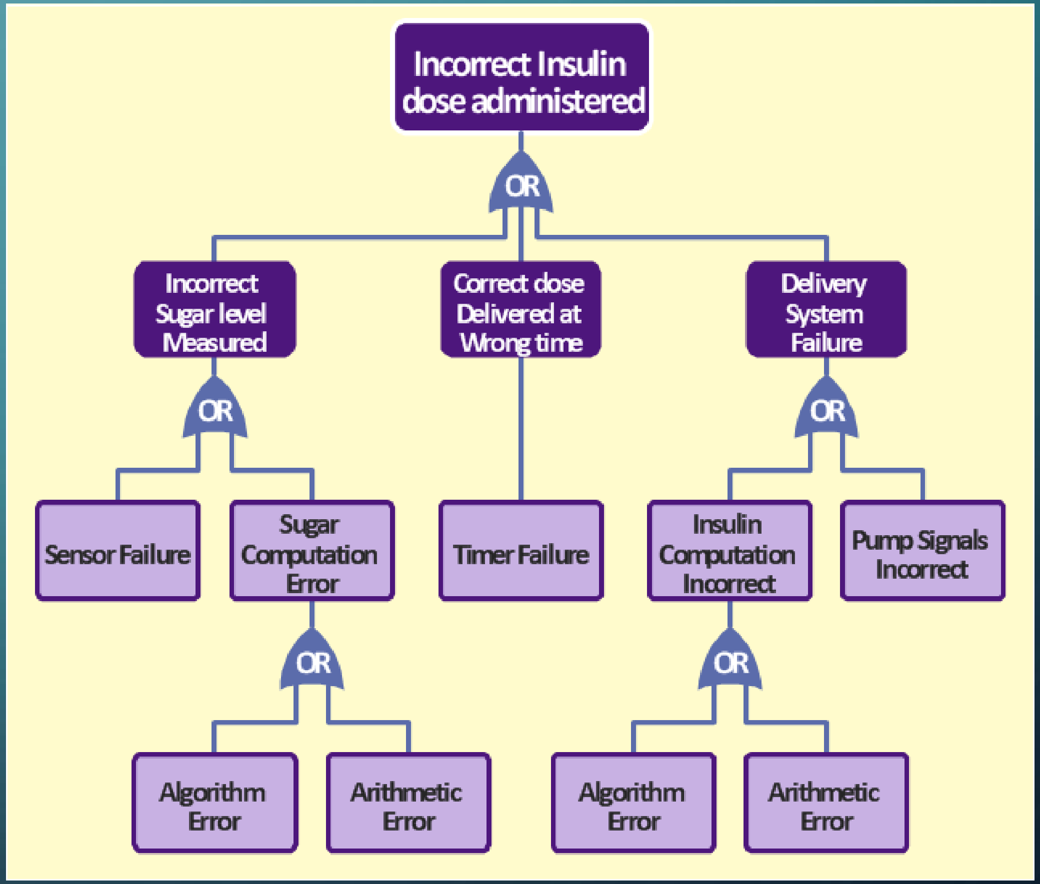
www.verifyrecruitment.com/blog/index.php/site-reliability-engineer/



www.spark.apache.org/docs/latest/cluster-overview.html

Traditional Approaches

- Fault trees
  - Minimal cut sets
  - Minimal path sets
- Gran and Helminen
  - Nuclear power plants
    - Software system
    - Plant hardware
- Halden project (HRP)
- Amasaki – software quality assessment
- Requires domain experts to
  - Evaluate prior probabilities
  - Understand the structure of BN.
  - Construct the BN



www.conceptdraw.com/a1579c3/preview

Challenges

- System domain specific
- Finding relations between components
- Systems evolve
- Finding domain expert
  - Difficult
  - Costly
  - Unintentional mistakes
  - Need for keeping same expert

www.technologynetworks.com/informatics/lists/7-data-challenges-in-the-life-sciences-288265

K2 Algorithm

Notations

- α_{ijk}
- α_{ij}
- π_i
- q_i
- d_i

- $P(A)$
- $P(A|B) = \frac{P(A,B)}{P(B)}$
- $P(A|B, C) = \frac{P(A,B,C)}{P(B,C)}$
- $P(A|B, C) = P(B, C|A) * \frac{P(A)}{P(B,C)}$

Pseudo Code

- Machine learning algorithm
- Scoring function
- Heuristic function
- Efficient(polynomial time) $O(n^2)$
- ~~System specific~~
- Historical data about the system

Algorithm $K2\ (T, u)$

Input: A dataset T of historical observations on system S, an upper bound u for the number of parents

Output: A full BN B.

1. For each column i in dataset T
   - Create node X_i and add it to B.
   - $\pi_i = \phi$ for node X_i.
   - Calculate $f(i, \pi_i)$ using empty set ϕ
   - While the size of $\pi_i \leqslant u$
     - Let X_z be a node preceding node X_i
     - Calculate $f(i, \pi_i\ U\ \{X_z\})$ using X_z
     - If the new score is better than the previous score
       - Add X_z to π_i permanently
2. Return B

$$f(i, \pi_i) = \prod_{j=1}^{q_i} \frac{(d_i - 1)!}{(\alpha_{ij} + d_i - 1)!} \prod_{k=1}^{d_i} \alpha_{ijk}!$$

Output depends on initial linear ordering of components.

Illustrative Step By Step Example

Dataset for example

observation	X_1	X_2	X_3	X_4	X_5	system behavior
1	1	1	0	0	0	0
2	0	1	1	1	1	1
3	1	1	0	1	1	1
4	0	0	0	0	0	0
5	1	0	1	0	0	1
6	0	0	0	0	0	0
7	1	1	0	1	1	1
8	0	1	1	1	1	1
9	1	0	1	0	0	1
0	0	0	0	0	0	0

f score for all possible candidate parent sets

parent set	f score
{}	1/2772
$\{X_1\}$	1/3600
$\{X_2\}$	1/3600
$\{X_1, X_2\}$	1/144

$$f(X_2, \phi) = \frac{(2-1)!}{(10+2-1)!} \prod_{k=1}^{2} \alpha_{i0k}! = \frac{1!}{11!} \times 5! \times 5! = \frac{1}{2772}$$

$$f(X_2, \{X_1\}) = \frac{(2-1)!}{(5+2-1)!} \prod_{k=1}^{2} \alpha_{21k}! \times \frac{(2-1)!}{(5+2-1)!} \prod_{k=1}^{2} \alpha_{22k}! = \frac{1}{3600}$$

Result Of K2 Algorithm

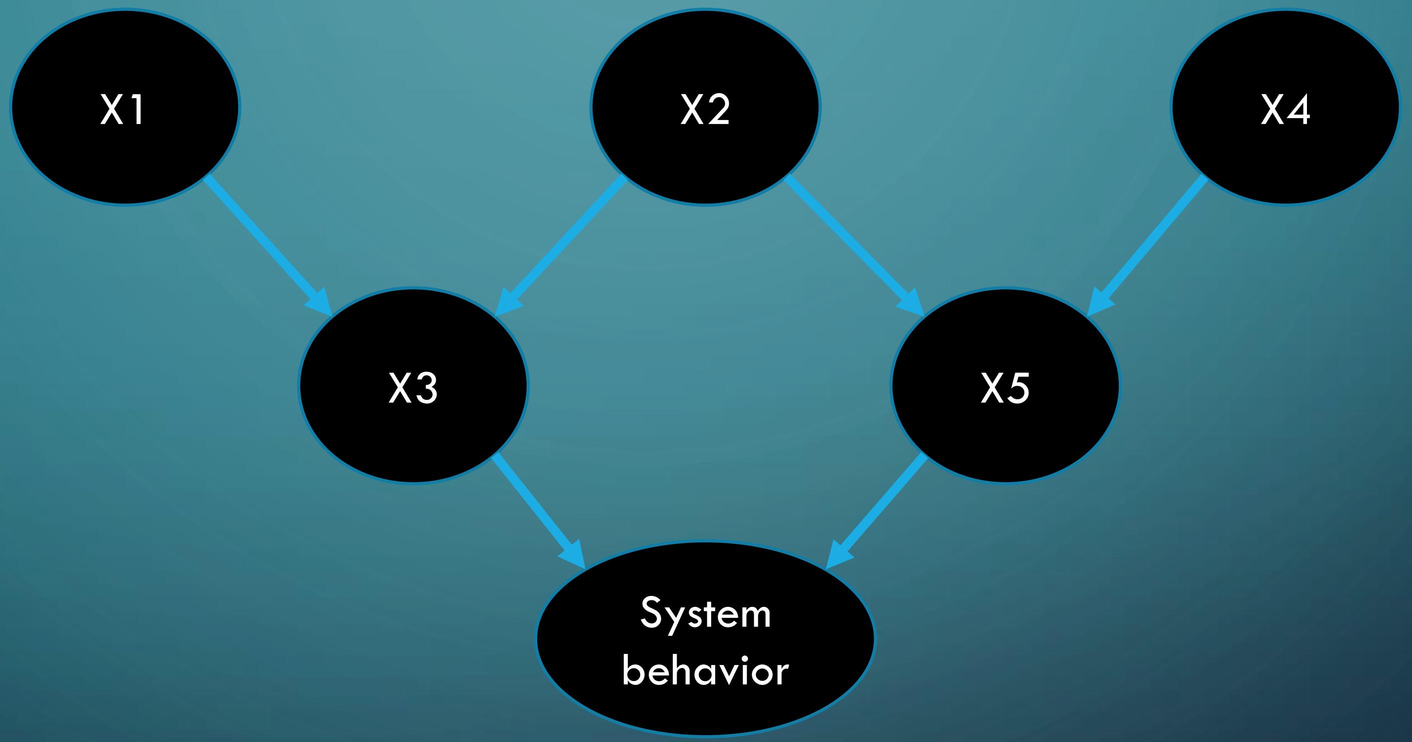


Fig. 1

CPT Of The X_3 Node

parents		probability	
X_1	X_2	$X_3 = 0$	$X_3 = 1$
0	0	1	0
0	1	0	1
1	0	0	1
1	1	1	0

Probability Of Success Calculation For Node X_3

$$p(X_3 = 1) = \sum p(X_1, X_2, X_3 = 1) =$$

$$p(X_1 = 0, X_2 = 0, X_3 = 1) + p(X_1 = 0, X_2 = 1, X_3 = 1)$$
$$+ p(X_1 = 1, X_2 = 0, X_3 = 1) + p(X_1 = 1, X_2 = 1, X_3 = 1)$$
$$= 0.5 * 0.5 * 0 + 0.5 * 0.5 * 1 + 0.5 * 0.5 * 1 + 0.5 * 0.5 * 0 = 0.5$$

CPT Of The System Behavior Node

parents		probability	
X_3	X_5	System Behavior = 0	System Behavior = 1
0	0	1	0
0	1	0	1
1	0	0	1
1	1	0	1

Experimental Analysis

How To Increase Performance ?
How To Reduce Error Rate ?

- More observations (more data) results more accuracy.
- More observations (more data) increases computation time.
- Trade OFF
- A_{FP}
- A_{FN}
- A_T
- $p = (A_{PF} + A_{PN}) / A_T$

Some BNs For Test K2 Algorithm

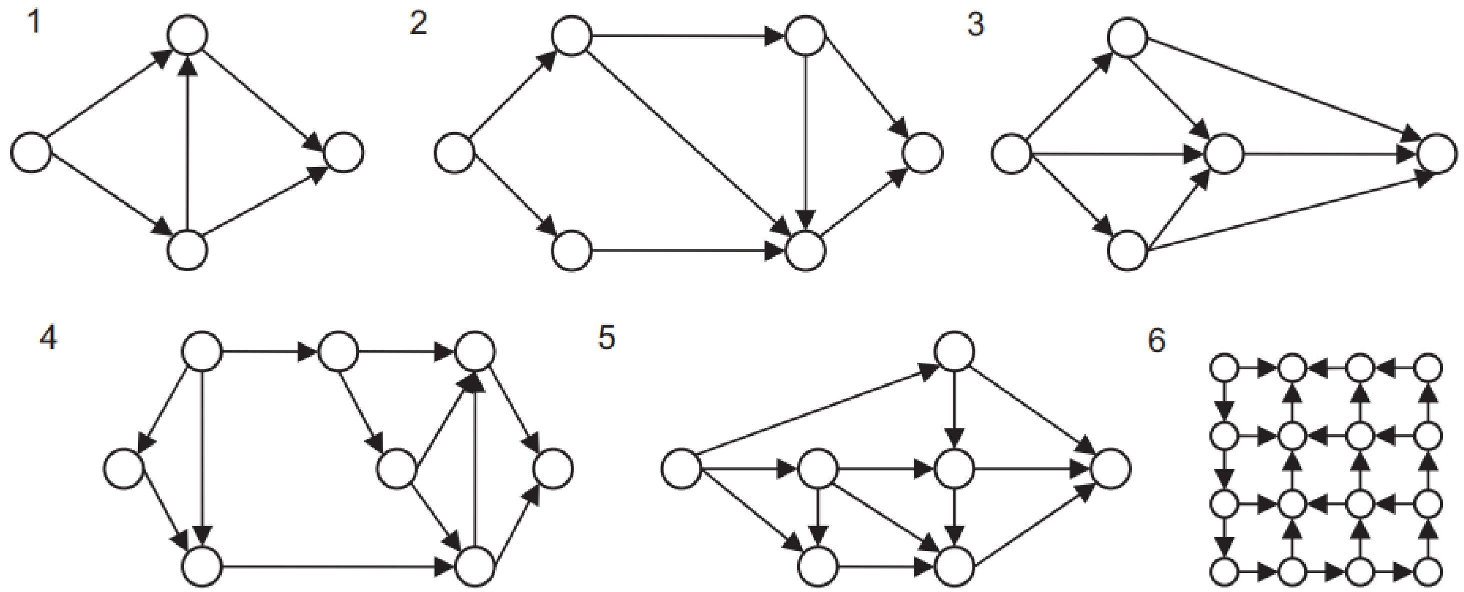


Fig. 2. Case BN tested on the *K2* algorithm.

Performance-Runtime Chart

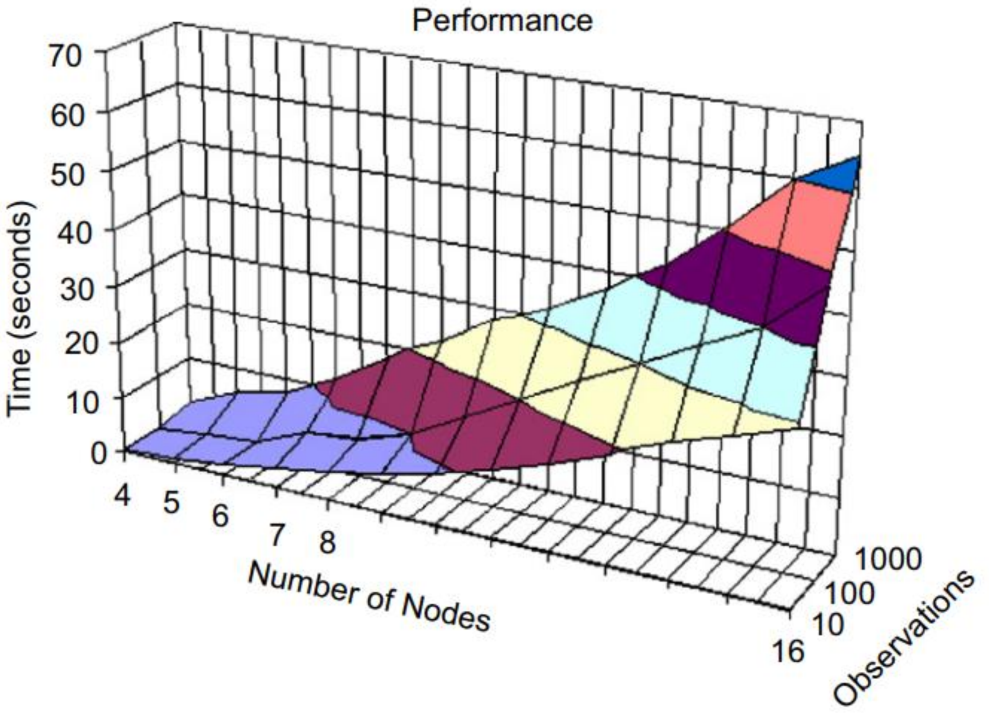


Fig. 3. Running time of the *K*2 algorithm.

Accuracy-Association Chart

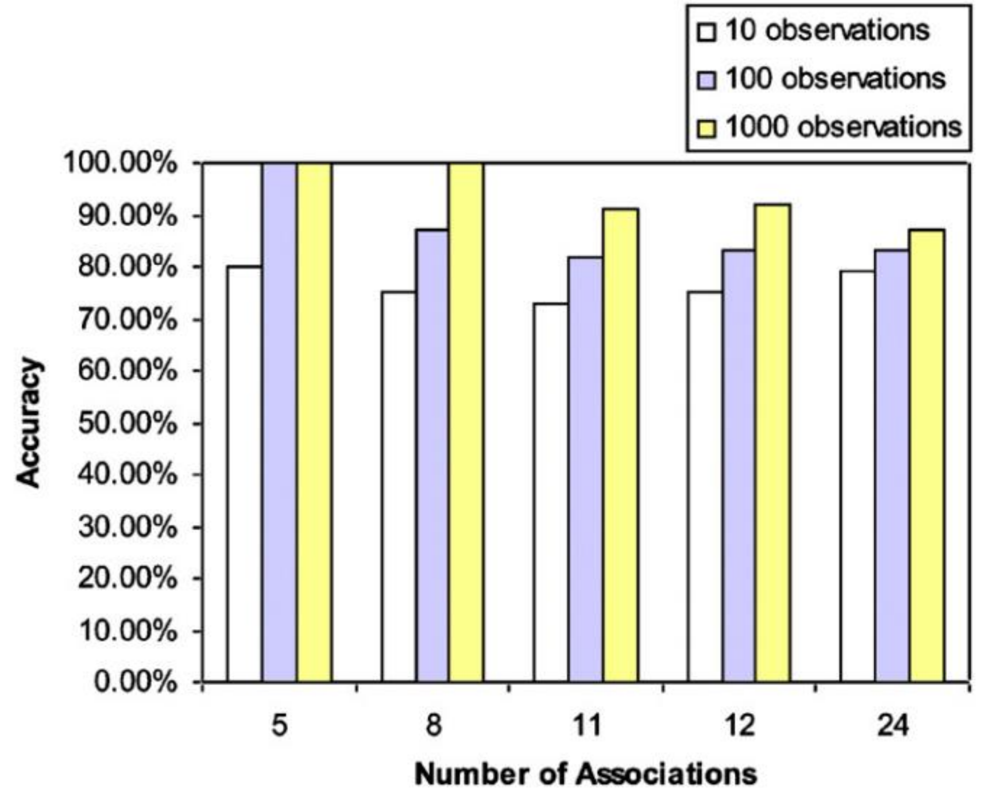


Fig. 4. Accuracy of results in BN construction.

Error Rate - Time Chart

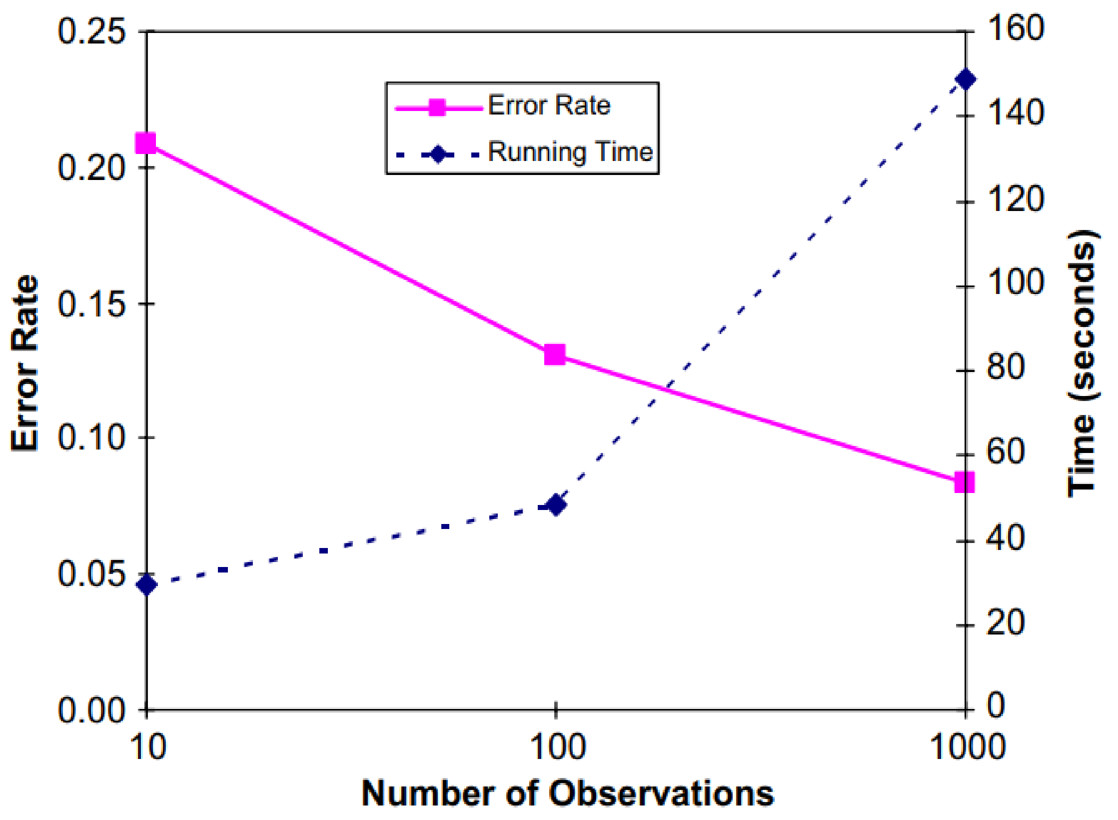


Fig. 5. Performance vs. accuracy curve for the *K*2 algorithm for case BN #6.

Conclusion

Conclusion

- Numerous method for estimating reliability but challenges still remained
- K2 algorithm:
  - No need to domain expert
  - But needs historical data about the system
  - Efficient for substantially large systems
- Can further be improved when the already existing associations between components are taken in to account(needs a domain expert of the system)

References

1. Ozge Doguc , Jose Emmanuel Ramirez-Marquez.A generic method for estimating system reliability using Bayesian networks.(February 2009).
2. Christopher M. Bishop.(2006).Pattern Recognition and Machine Learning.
3. Stuart J. Russell and Peter Norvig. Artificial Intelligence A Modern Approach 3th Edition.
4. Helge Langseth , Luigi Portinale. Bayesian networks in reliability.(19 January 2006).
5. Sankaran Mahadevan ,Ruoxue Zhang , Natasha Smith.Bayesian networks for system reliability reassessment.(2001).

?

THANK YOU