Evolved Bayesian Network Models of Rig Operations in the Gulf of Mexico

François A. Fournier, John McCall, Andrei Petrovski, Peter J. Barclay

Robert Gordon University, Aberdeen - ODS-Petrodata Ltd.

Knowledge Transfer Partnerships

Content

- Drilling Rigs Operations
- Applications
- Data
- Bayesian Networks
- Structure learning algorithms
- Experiments
- Results
- Future work

Drilling Rigs Operations

- Exploration and Exploitation
- $400K to $600K per day
- Weeks/months-long operations
- Dynamic, highly competitive, and regionally-specific markets
- Drilling Rig Efficiency standards
- Empirical observations and experience

Drilling Rigs Operations

- "successful operation depends on many factors which are difficult to measure"
  Harris (1989)

- Typically, three main criteria for selecting drilling rig for hire: Osmundsen (2008)
  - technical suitability
  - price
  - availability

Objectives

- Drilling Rig selection
  - Select rig for particular demand
  - Recommend better rig to the user based on personal profile
- Rig Performance Forecasting
  - Performance expectation
  - Hire decision guidance
- Rig Scheduling
  - Completion time forecasting
  - Automatic scheduling advice

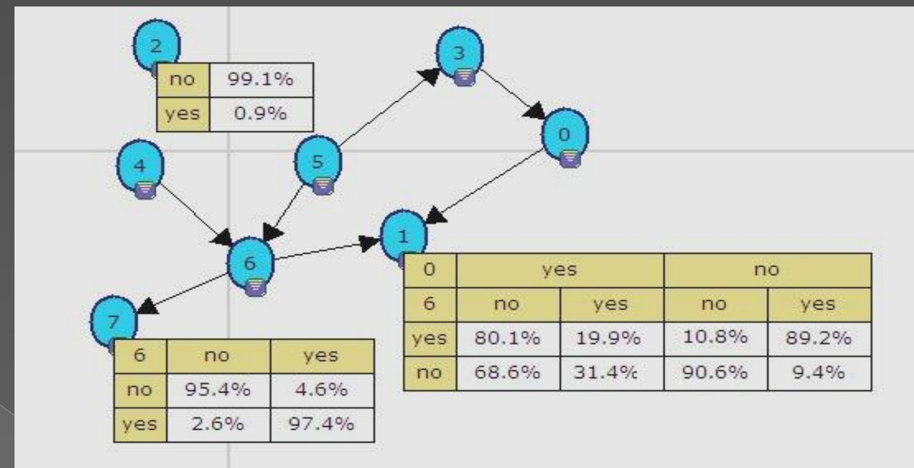
Data

- ODS-Petrodata: 25 years historical data
  - Rigs, Wells, Demands
  - Vessels, Seismic, Renewables
- Wells-Rigs-Demands Dataset (WRD)
  - Potentially 700-1000 factors
  - Pilot experiments: 17 key factors
  - Initial data: 6670 items

Field Name	# of values
Well Phase	6
Well Deviated	4
Well Type	6
Well Status	7
Well Result	17
Days On Location (d)	11
Number of Days to Total Depth (d)	10
Total Vertical Depth (d)	18
Total Footage Drilled (d)	18
Average Feet drilled Per Day (d)	16
Shore Base	54
Region	59
Water Depth (d)	10
Rig Type	6
Harsh Environment Capability	2
Rig Owner	72
Rig Contractor	70

Bayesian Networks

- Directed Acyclic Graph
- Conditional Probabilities Table



- Network constructed for a given problem
- Learning the structure is NP hard
- Number of possible structures grows super-exponentially with the number of variables

K2

Cooper and Herskovitz (1992)

- Greedy search
- Search Space of ordering
- Incrementally add links
- Maximise score
- Stops when no more additions improve the score
- Prone to local optima
- Sensitive to node ordering

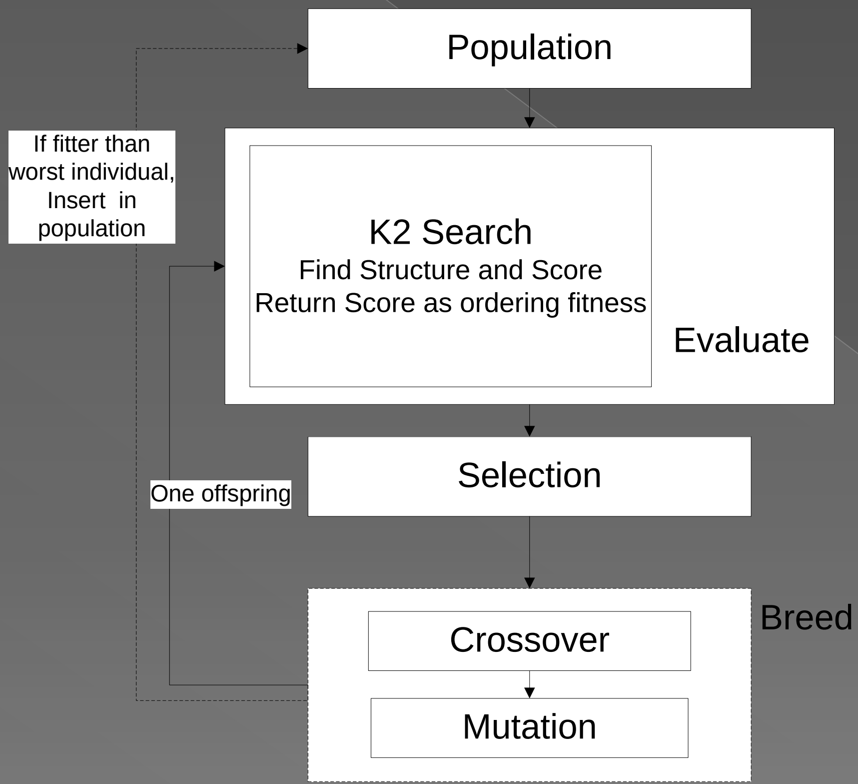
K2-CH score

- Requires exploring all combinations of values throughout the dataset
- Product of factors
- Main computational effort of the algorithm

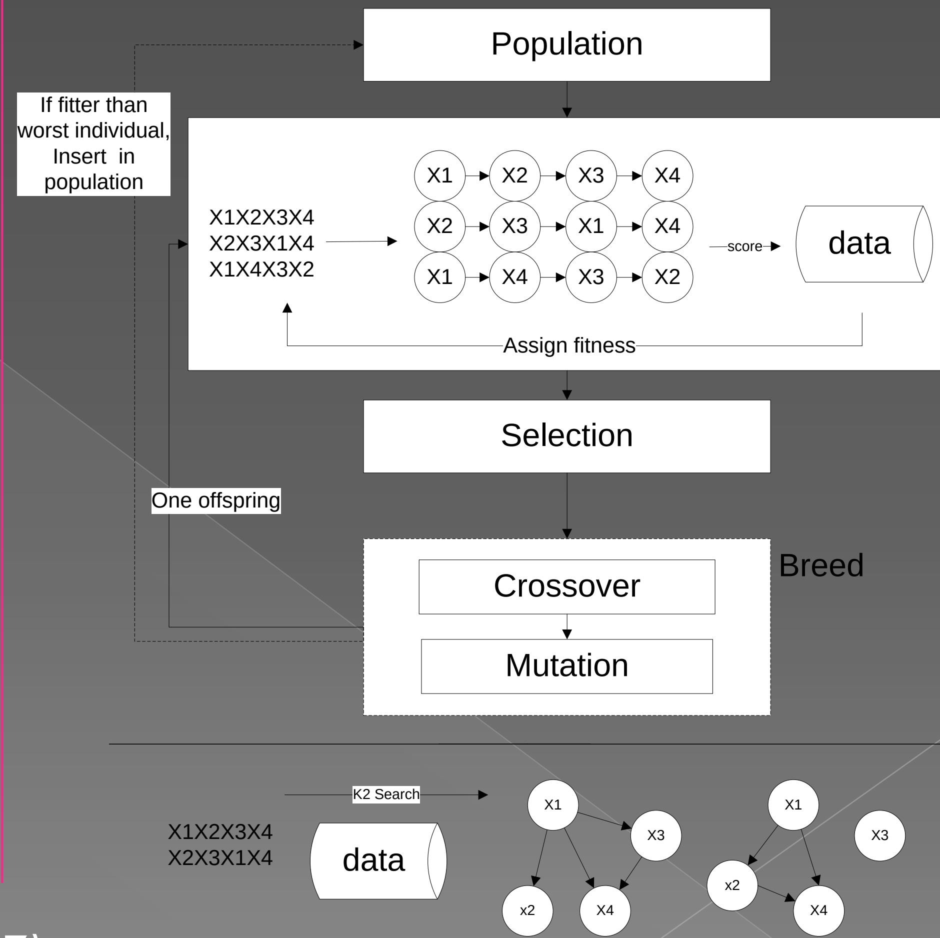
$$P(B_s, D) = P(B_s) \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}!$$

K2GA

ChainGA



a) K2GA



b) ChainGA

Larrañga et al. (1996), Kabli et al. (2007)

Experiments

- 17 nodes
- 2500 cases
- K2GA and ChainGA
- 45 runs each
- 200 generations
- 30 node orderings population size
- 0.05 Mutation rate
- 0.9 Crossover rate
- Tournament selection size 4

Size dataset	K2GA	ChainGA
100	20 mins	3 mins
2500	42 hours	13 hours

Results

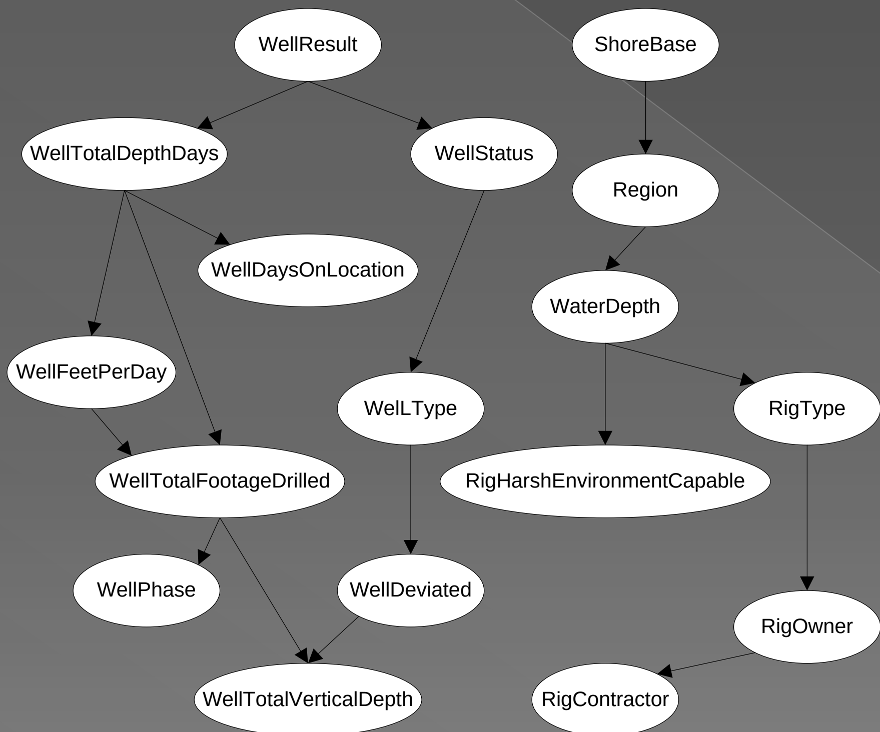
2-tailed t-test of Best Individuals K2 Score across all runs

	N	Mean Score	Standard Deviation	P
K2GA	45	-56197.44	205.2	< 0.0005
ChainGA	45	-66434.34	1237.7	< 0.0005

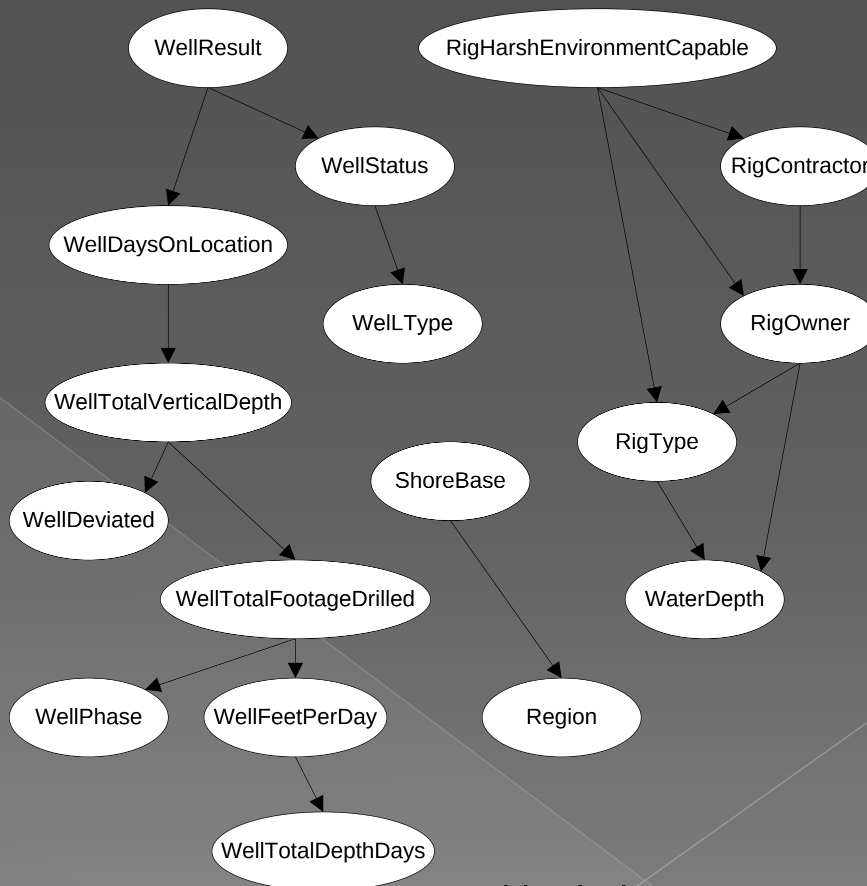
Time per run over all runs

	Average	Standard Deviation
K2GA	42h 28min	5h 9min
ChainGA	11h 1min	1h 11min

Best-scoring structures



a) K2GA



b) ChainGA

Future work

- Different algorithms (ACO, PSO)
- Different scoring metrics (MDL, Bdeu)
- More data
- More variables (100+)
- Larger geographical regions

Conclusion

- New problem
- Suitable algorithms for this dataset
- Structures approved by experts
- Trade-off between results and computation time
- Support of decision making
- Real industry data

Evolved Bayesian Network Models of Rig Operations in the Gulf of Mexico

François A. Fournier, John McCall,
Andrei Petrovski, Peter J. Barclay

Robert Gordon University, Aberdeen - ODS-Petrodata Ltd.

Knowledge
Transfer
Partnerships

ODS-PETRODATA