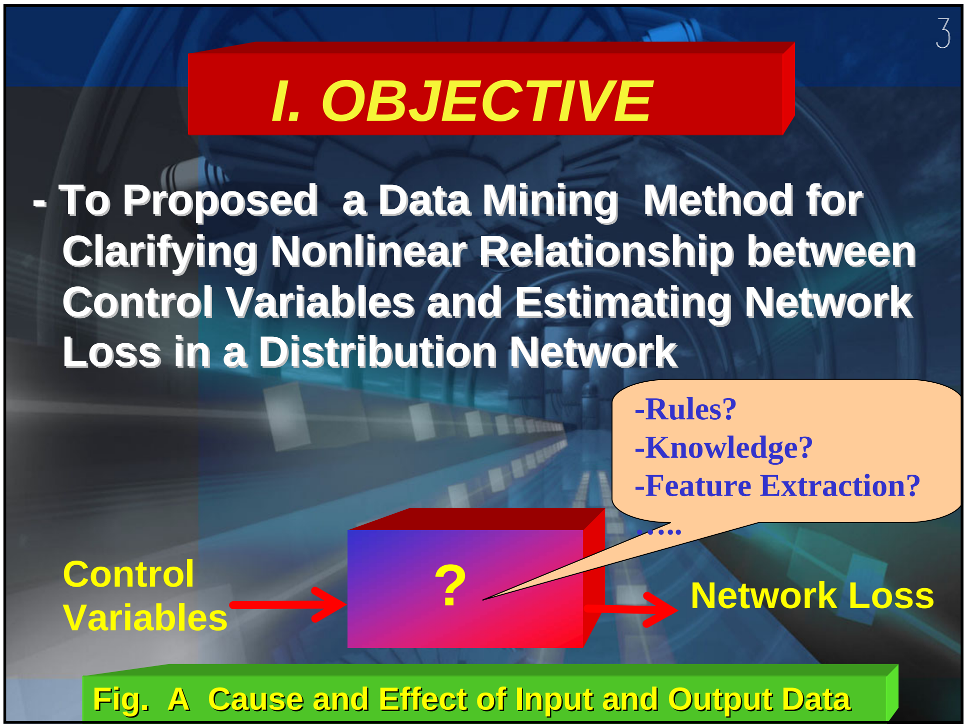
IEEE T&D2008 Panel

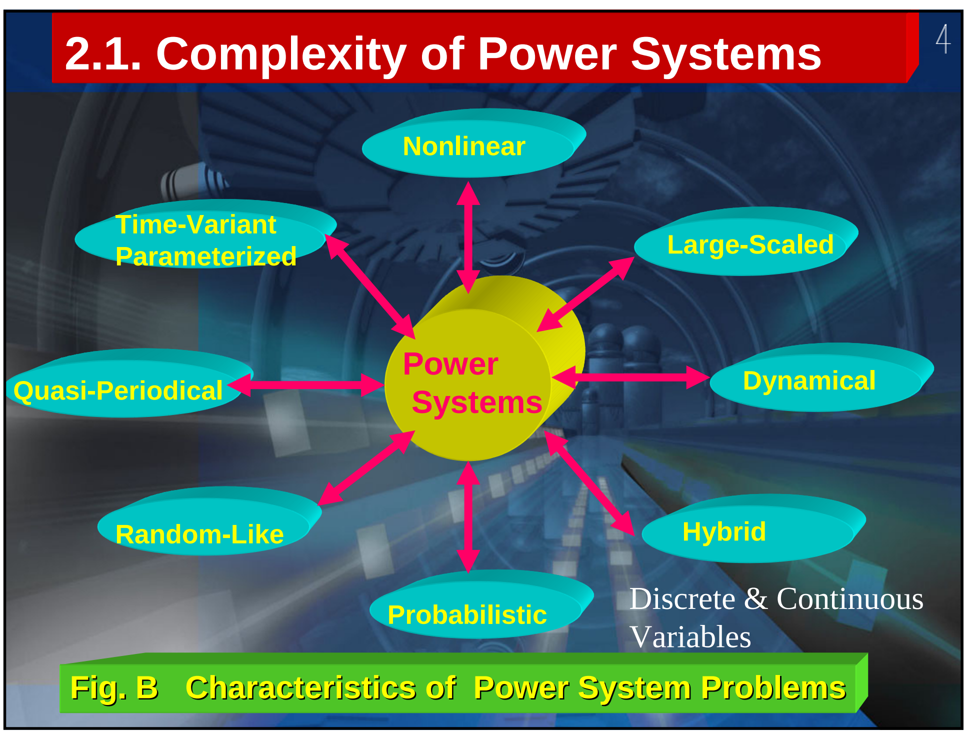
A Data Mining Technique for Three-Phase Distribution Network Voltage Control

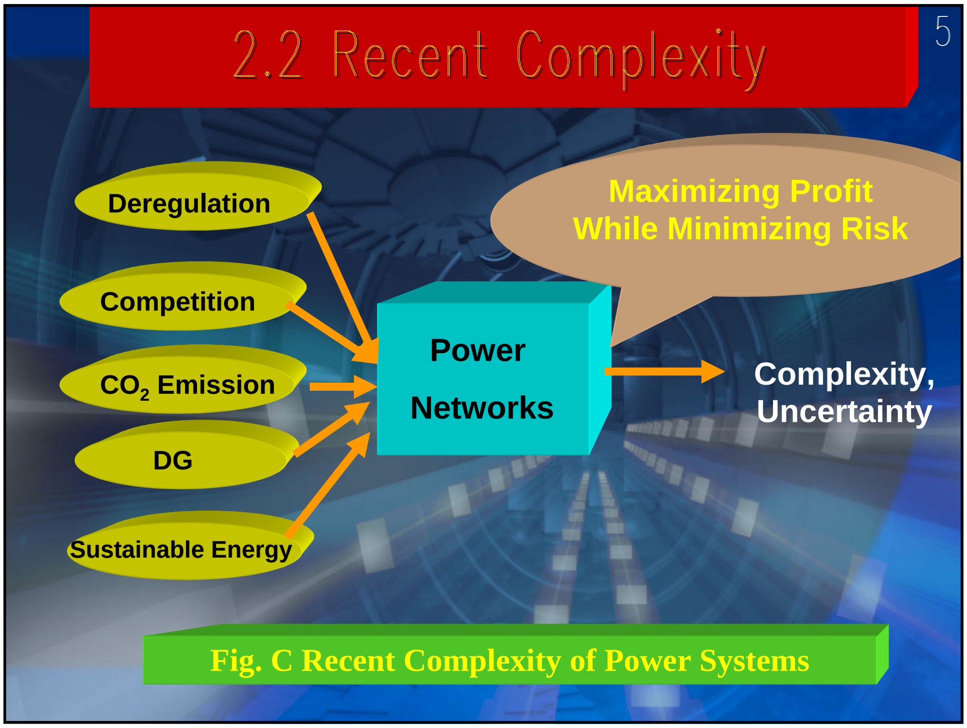
H. Mori, K. Shimomugi and Y. Umezawa
Dept. of Electronics・Bioinformatics
Meiji University
Tama-ku, Kawasaki 214-8571
Japan
hmori@isc.meiji.ac.jp

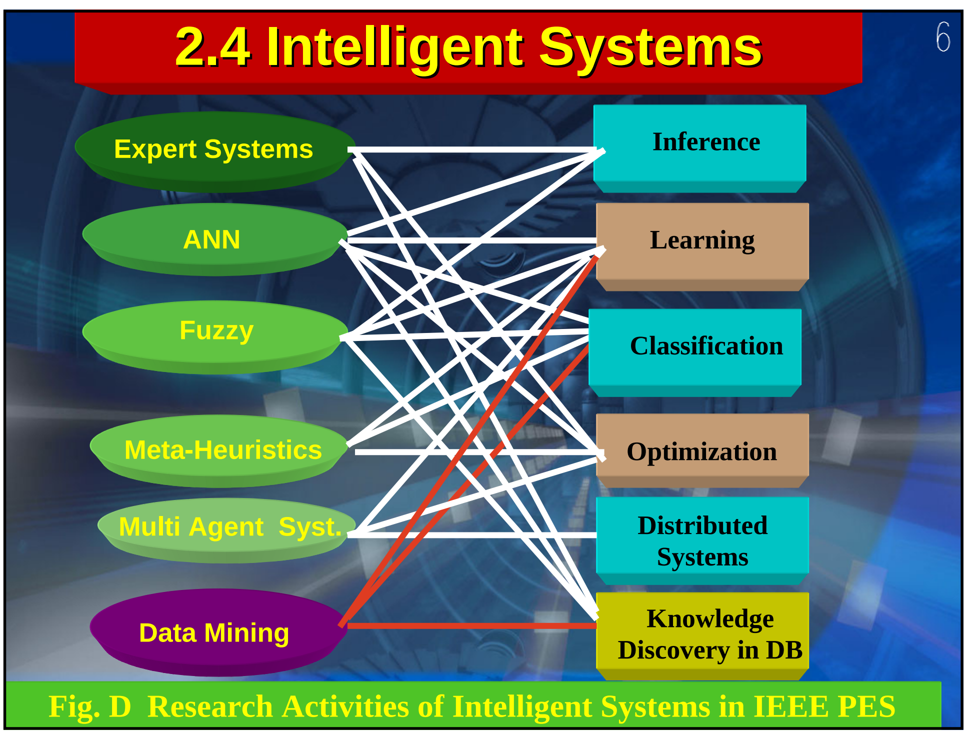
OUTLINE

I. Objective
II. Background
III. Three-Phase Distribution Power Flow
IV. Proposed Method
V. Simulation
VI. Conclusion


3
I. OBJECTIVE
- To Proposed a Data Mining Method for Clarifying Nonlinear Relationship between Control Variables and Estimating Network Loss in a Distribution Network
-Rules?
-Knowledge?
-Feature Extraction?
…..
Control Variables
?
Network Loss
Fig. A Cause and Effect of Input and Output Data


4
2.1. Complexity of Power Systems
Nonlinear
Time-Variant Parameterized
Large-Scaled
Power Systems
Quasi-Periodical
Dynamical
Random-Like
Hybrid
Probabilistic
Discrete & Continuous Variables
Fig. B Characteristics of Power System Problems


2.2 Recent Complexity
5
Deregulation
Competition
CO_2 Emission
DG
Sustainable Energy
Power
Networks
Maximizing Profit
While Minimizing Risk
Complexity,
Uncertainty
Fig. C Recent Complexity of Power Systems


2.4 Intelligent Systems
6
Expert Systems
ANN
Fuzzy
Meta-Heuristics
Multi Agent Syst.
Data Mining
Inference
Learning
Classification
Optimization
Distributed
Systems
Knowledge
Discovery in DB
Fig. D Research Activities of Intelligent Systems in IEEE PES

7

2.5 Roles of Data Mining

- To Understand Complicated Data with Some Rules
- To Extract Important Features That are Known and/or Unknown
- To Construct More Reasonable Models/Strategies

8

2.6 What PS Areas Is Data Mining Applied to?

- Dynamic Security Assessment (Wehenkel, *et al.*, '94; Rovnyak and Thorp, '94)
- Load Forecasting (Mori & Kosemura,'01)
- Power System Control Center (Lambert-Torres, *et al.*,'02)
- Data Profiling of Customers (Kitamura, *et al.*,'02) *etc.*

9

2.. Distribution Automation

- Optimal network reconfiguration[2,6]
- Network loss minimization[3]
- Voltage and reactive power control[1,5,10]
- Distribution network service restoration[4,9]
- State estimation
- Load estimation[7,8]

10

2.. Three-Phase Distribution Network Voltage Regulation

- To Maintain Nodal Voltage. with Input Variables Such as Voltage Regulators.

- However, It Is Hard to Determine the Appropriate Tap Positions Because the Tap Changes Bring about the Voltage Violation at the Secondary Side of the Regulator due to Discrete Variable Adjustment.

11

2.. Three-Phase Distribution Network Voltage Regulation.Continued)

.

- A Set of Power Flow Calculations Are Needed to Maintain the Nodal Voltage within the Upper and the Lower Bounds.
- The Trial-and-Error Method is Necessary to Evaluate Solutions.
- Therefore, the Three-Phase Distribution Network Voltage Control Needs a Lot of Computational Effort to Maintain the Voltage Profile.

12

III. Three-Phase Distribution Power Flow

3.1 Concept of Three-Phase Backward-Forward Sweep Method.

- Backward-Sweep is to Add up the Injection Currents from the Ending Node to the Distribution Substation.
- Forward-Sweep is to Calculate the Nodal Voltage from the Substation to the Ending Node.
- The Nodal Voltages are Evaluated by Repeating the Above Calculation Until the Convergence Criterion Is Satisfied.

3.2 Algorithm

13

Step 1: Set the initial conditions of the nodal voltages.

Step 2: Evaluate the injection current at each node by the following equation:

.

$$\begin{bmatrix} I_{ia} \\ I_{ib} \\ I_{ic} \end{bmatrix}^{(k)} = \begin{bmatrix} \left(S_{ia}/V_{ia}^{(k-1)}\right)^* \\ \left(S_{ib}/V_{ib}^{(k-1)}\right)^* \\ \left(S_{ic}/V_{ic}^{(k-1)}\right)^* \end{bmatrix} - \begin{bmatrix} Y_{iaa}^* & Y_{iab}^* & Y_{iac}^* \\ Y_{iba}^* & Y_{ibb}^* & Y_{ibc}^* \\ Y_{ica}^* & Y_{icb}^* & Y_{icc}^* \end{bmatrix} \begin{bmatrix} V_{ia} \\ V_{ib} \\ V_{ic} \end{bmatrix}^{(k-1)} \quad (1)$$

Step 3: Add up the node injection currents from the ending node to the substation.

$$\begin{bmatrix} J_{la} \\ J_{lb} \\ J_{lc} \end{bmatrix}^{(k)} = -\begin{bmatrix} I_{ja} \\ I_{jb} \\ I_{jc} \end{bmatrix}^{(k)} + \sum_{m \in M} \begin{bmatrix} J_{ma} \\ J_{mb} \\ J_{mc} \end{bmatrix}^{(k)} \quad (2)$$

3.2 Algorithm(Continued)

14

Step 4: Calculate the nodal voltage from the substation to the ending node.

$$\begin{bmatrix} V_{ja} \\ V_{jb} \\ V_{jc} \end{bmatrix}^{(k)} = \begin{bmatrix} V_{ia} \\ V_{ib} \\ V_{ic} \end{bmatrix}^{(k)} - \begin{bmatrix} z_{aa,l} & z_{ab,l} & z_{ac,l} \\ z_{ba,l} & z_{bb,l} & z_{bc,l} \\ z_{ca,l} & z_{cb,l} & z_{cc,l} \end{bmatrix} \begin{bmatrix} J_{la} \\ J_{lb} \\ J_{lc} \end{bmatrix}^{(k)} \quad (3)$$

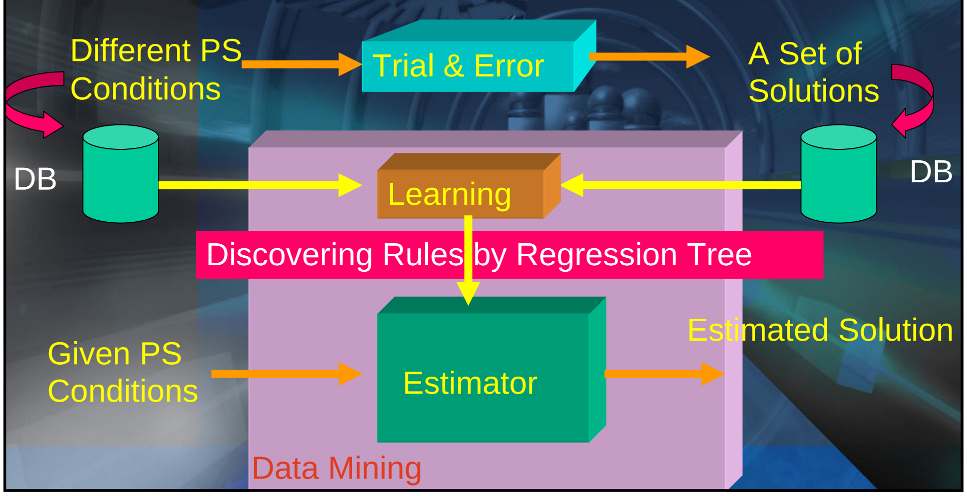
Step 5: Stop if the convergence criterion is satisfied. Otherwise, return to Step 2.

Open Problems in Three-Phase Distribution Voltage Control

- *The Power Flow Calculation Does Not Converge to a Solution in a Certain Power System Conditions.*
- *There are Some Cases that Voltage Regulator Tap Changes Are Not Appropriate Since the Change Triggers the Voltage Violation at the Neighbor Nodes .*

IV. Proposed Method

4.1 Outline of Proposed Method



17

4. 2 Outline of DM

Data Mining

To Discover Important Rules in Large Data Base

Data Mining Methods

- Pattern Recognition (ANN)
- Fuzzy Theory
- Decision Tree, *etc.*

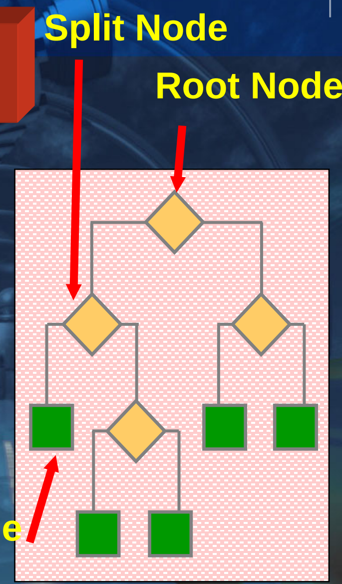


Fig. E Decision Tree

18

Table A Difference between Classification and Regression Trees

Decision Tree	Output	Conventional Methods
Classification	Qualitative	CART, ID3 C4.5
Regression	Quantitative	CART

4.3 Regression Tree

The algorithm of the regression tree consists of the following three steps:

(A1)

Step 1: Grow a tree.

Step 2: Carry out pruning the tree.

Step 3: Select the optimal tree in candidates.

4.3 Regression Tree (Continued)

- **Growth**

  Minimization of Error After Splitting

$$R(n) = \frac{V(n)}{V_0}$$

(A1)

$R(n)$: Error of Node n

$V(n)$: Variance of Learning Data Belonging to Node n

V_0: Variance of All Learning Data

- **Pruning**

  Simple Structure of Regression Tree

- **Error Estimate**

  Cross-Validation Method

Constructing the Tree

Error Reduction

$$\Delta R(s,t) = R(t) - R(t_L) - R(t_R) \qquad (5)$$

Where, $R(s,t)$: Degree of Error Reduction in Case Where Attribute s at Node t. s: Attribute, t: Parent Node, $R(t)$: Sum of Squared Error of Parent Node. $R(t_{L(R)})$: Error of Left-side (Right-side) Child Node

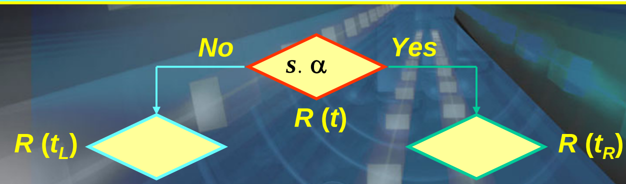


Fig. F Process of Constructing Tree

Pruning

$$r(t_p)=r^{CV}(t_p)+\sigma(t_p) \qquad (2)$$

Where, r: Error. $r^{CV}(.)$: Cross-Validation Error. $\sigma(.)$: Standard Deviation of Cross-Validation Error. t_p : Pruned Tree Number

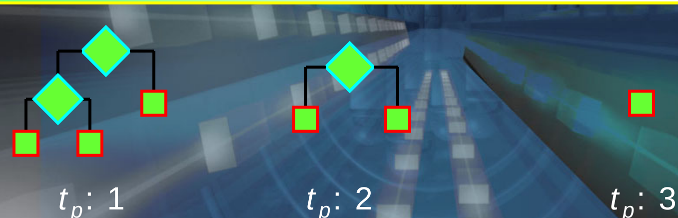


Fig. G Pruning Process

4.4 Addition Information of Regression Tree

The variable importance shows the rate of the error reduction to the whole. It may be written as :

We Can Identify Important variables by the Index

$$VI(x) = \frac{\sum_{t \in \tilde{N}_s} i(x,t)}{\max_{x \in X} \left(\sum_{t \in \tilde{N}_s} i(x,t) \right)} * 100 \quad (9)$$

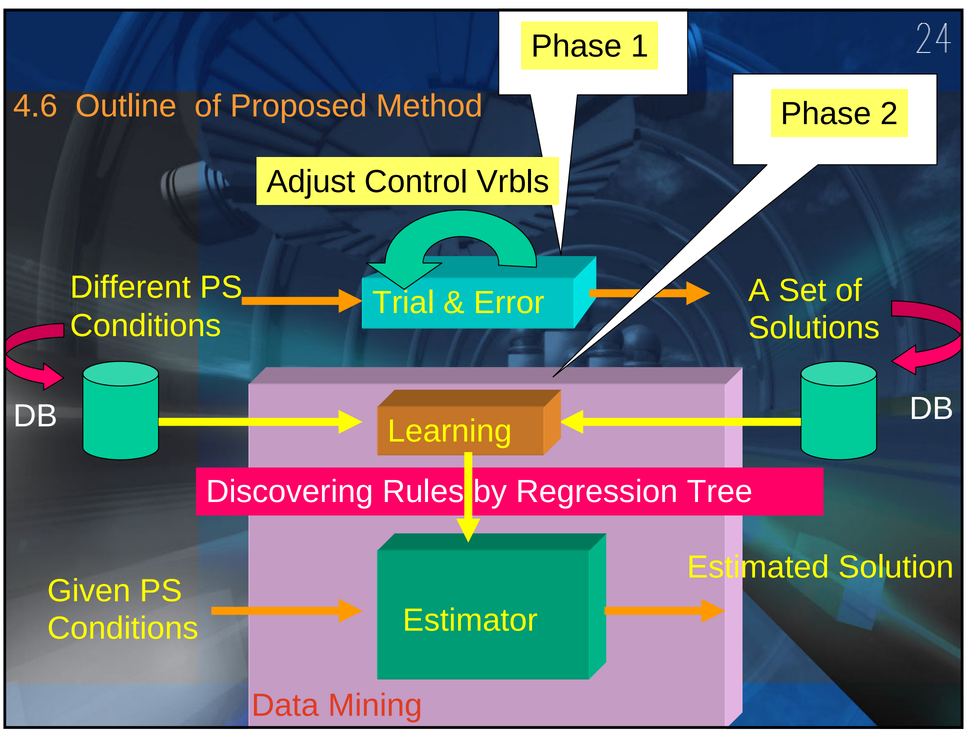
where

VI(x): variable importance of input variable *x*

X: set of input variables

Ns: set of split nodes

i(x,t): degree of improvement of input variable x at node i



25

4.7 Algorithm of Proposed Method

The Target is to Evaluate a Feasible Three-Phase Power Flow Solution That Meets the Upper and Lower Bounds of Nodal Voltages:

(A1)

Phase 1(Creation of Learning Data)

Step 1: Set initial conditions for each nodal voltage.

Step 2: Compute the injection currents at each node

Step 3: Add up all the currents from the ending node to the substation.

26

4.7 Algorithm of Proposed Method (Continued)

Step 4: Calculate the nodal voltage from the substation to the ending node.

Step 5: Evaluate the power mismatch.

(A1)

Step 6: Stop if the converged solution within the lower and upper bounds is obtained. Otherwise, adjust the voltage regulators and return to Step 2.

Phase 2(Classify Accumulated Data into Terminal Nodes with the Regression Tree)

27

V. SIMULATION

Sample System: Modified IEEE Three-Phase 34-node Distribution System

Modification:

- Branch between Nodes 882 and 888 has 100ft. Feeder Rather Than the Transformer.
- There are two voltage regulators .One tap raises 0.75V in 120V base.
- Nodes 844 and 848 have Three Kinds of Capacitor banks: 50kVar, 100kVar, and 150kVar
- Two Loads at Nodes 844 and 890 are Randomly Changed(Gaussian , original data ±20%)

# of PS Conditions : 500

28

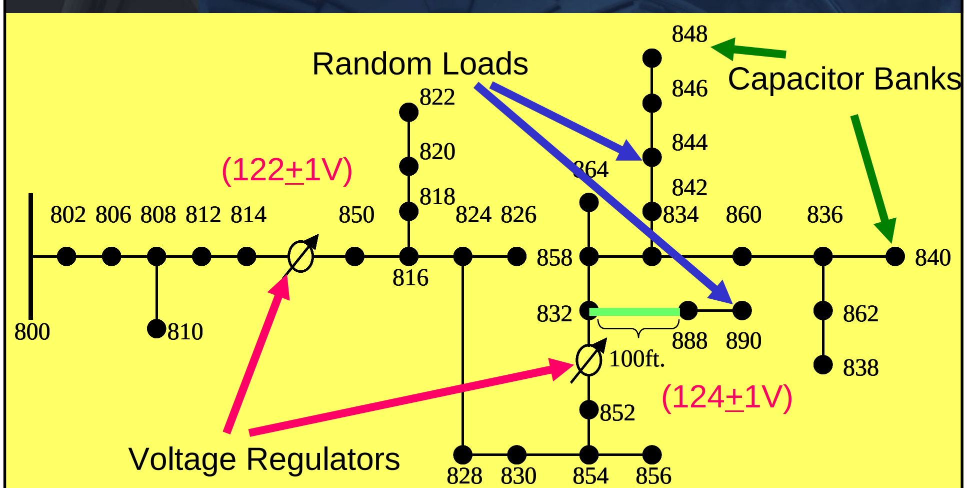


Fig. 1 IEEE 34-node System

V. SIMULATION

Input Variables

- **YPL(1~3): three-phase active load at node 844**
- **YQL(1~3): three-phase reactive load at node 844**
- **DPL(1~3): three-phase active load at node 890**
- **DQL(1~3): three-phase reactive load at node 890**
- **TAP(1~3): voltage regulator taps 1~6**
- **CAP(1~6): capacitor banks 1~6**

Output Variable

- Active Power Network Loss

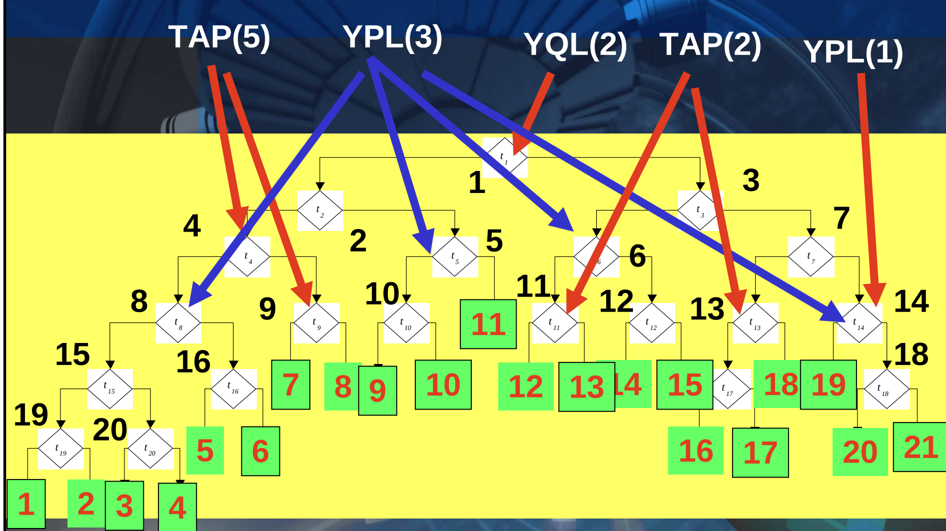


Fig. 2 Constructed Tree

31

Table 1 Obtained Rule at Each Splitting Node

t	Obtained Rules	*t*	Obtained Rules
1	**YQL(2) 145779.453**	11	**TAP(2) 3.500**
2	**TAP(1) 10.500**	12	**TAP(1) 10.500**
3	**DQL(2) 147600.250**	13	**TAP(2) 4.500**
4	**TAP(5) 12.500**	14	**YPL(1) 137917.828**
5	**YPL(3) 151518.078**	15	**DPL(1) 150528.922**
6	**YPL(3) 146388.703**	16	**TAP(4) 11.500**
7	**TAP(4) 11.500**	17	**DPL(3) 148899.250**
8	**YPL(3) 153054.656**	18	**YPL(3) 157364.531**
9	**TAP(5) 13.500**	19	**DPL(3) 153884.625**
10	**DPL(1) 146224.453**	20	**DPL(3) 148228.406**

32

Table 1A Obtained Rule at Terminal Node 11

t	Obtained Rules	*t*	Obtained Rules
1	**YQL(2) 145779.453**	11	**TAP(2) 3.500**
2	**TAP(1) 10.500**	12	**TAP(1) 10.500**
3	**DQL(2) 147600.250**	13	**TAP(2) 4.500**
4	**TAP(5) 12.500**	14	**YPL(1) 137917.828**
5	**YPL(3) 151518.078**	15	**DPL(1) 150528.922**
6	**YPL(3) 146388.703**	16	**TAP(4) 11.500**
7	**TAP(4) 11.500**	17	**DPL(3) 148899.250**
8	**YPL(3) 153054.656**	18	**YPL(3) 157364.531**
9	**TAP(5) 13.500**	19	**DPL(3) 153884.625**
10	**DPL(1) 146224.453**	20	**DPL(3) 148228.406**

IF YQL(2).145779.453, TAP(1).10.500, and YPL(3).151518.078, Then Network Loss is 285.186KW.

33

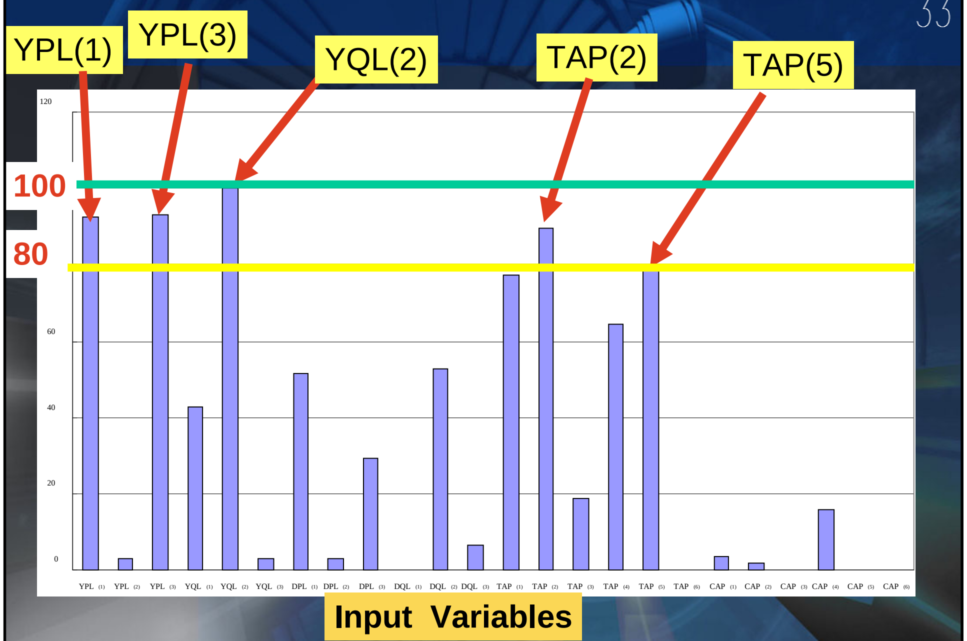


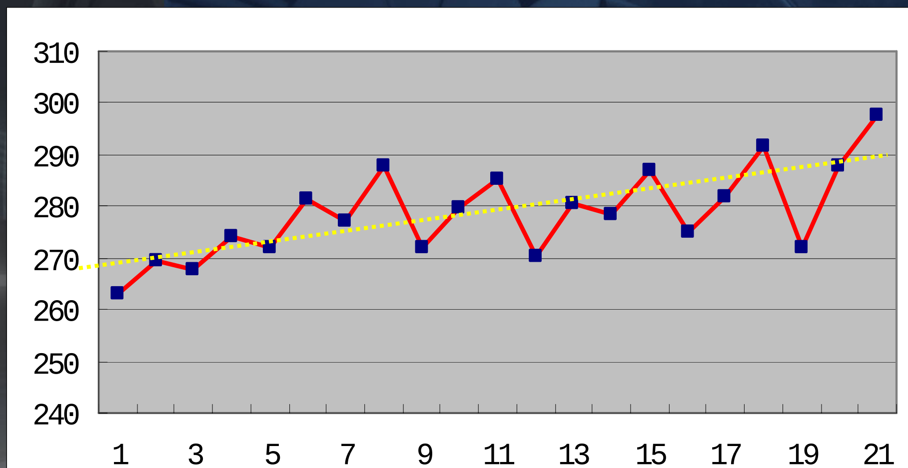
Fig. 3 Variable Importance for Input Variables

34

Table 2 Average Loss and No. of Data at Each Terminal Node

N_T	Ave. [kW]	No. of data	N_T	Ave. [kW]	No. of data
1	262.941	47	11	285.186	29
2	269.377	31	12	270.211	32
3	267.596	24	13	280.537	10
4	274.273	23	14	278.472	30
5	272.198	18	15	286.787	11
6	281.299	11	16	275.127	29
7	277.115	36	17	282.001	28
8	287.91	6	18	291.63	8
9	272.211	20	19	271.852	23
10	279.625	32	20	287.884	44
			21	297.566	8

Network Loss



Input Variable Number

Fig. 4 Relationship between Node # and Average Active Power Loss at Terminal Nodes

VI. CONCLUSION

- This paper has proposed a data mining technique for distribution network voltage regulation. The proposed method efficiently extracts rules for dealing with distribution network loss minimization. This paper presented a regression-tree-based method that clarifies the nonlinear relationship between control variables and the network loss in a three-phase distribution network.
- The proposed method was applied to the IEEE 34-node. A regression tree is constructed to find out rules in a distribution network voltage regulation. The simulation results have shown that a set of network conditions are optimally classified into the terminal nodes with if-then rules and the importance of input variable was clarified by the index That allows distribution network operators to understand the network conditions

37

References

- H. Mori and N. Kosemura, "Optimal Regression Tree Based Rule Discovery for Short-term Load Forecasting," Proc. of 2001 IEEE PES Winter Meeting, Vol. 2, pp.421-426, Columbus, USA, Jan. 2001.
- H. Mori, N. Kosemura, K. Ishiguro and T. Kondo, "Short-term Load Forecasting with Fuzzy Regression Tree in Power Systems," Proc. of 2001 IEEE International Conference on Systems, Man & Cybernetics, pp. 1948-1953, Tucson, AZ, U.S.A., Oct. 2001.
- H. Mori, N. Kosemura, T. Kondo and K. Numa ,"Data Mining for Short-term Load Forecasting," Proc. of 2002 IEEE PES Winter Meeting, Vol. 1, pp.623-624, New York, NY, USA, Jan. 2002.
- H. Mori and Y. Sakatani, "An Integrated Method of Fuzzy Data Mining and Fuzzy Inference for Short-term Load Forecasting," Proc. of ISAP (CD-ROM), Limnos, Greece, Aug. 2003.
- H. Mori, Y. Sakatani, T. Fujino and K. Numa, "An Efficient Hybrid Method of Regression Tree and Fuzzy Inference for Short-term Load Forecasting in Electric Power Systems," A..Lofti and M. J. Garibaldi (Eds.), “Applications and Science in Soft Computing,” pp. 287-294, Springer, Berlin, Germany, Nov. 2003.

38

2.3 Trends on Intelligent Systems in IEEE Power Engineering Society

- Expert Systems(1980~)
- Artificial Neural Net(1987~)
- Fuzzy Inference(1990~)
- Evolutionary Computation, Modern Heuristics or Meta-Heuristics(1990~)
- Multi-Agent Systems(1995~)
- Data Mining(2000~)