

シミュレーテッドアニーリングを用いた

5 J-10 定量的知識から定性的知識への写像関数最適化手法の評価

松本 一則

橋本 和夫

小花 貞夫

国際電信電話株式会社

1. はじめに

国際電話網の疎通悪化診断を行うエキスパートシステム(ES)^[1]では、「もし回線使用率が普通で、話中信号検出率が徐々に増加した場合、相手国の国内網の輻輳が考えられる」のように、専門家の知識を定性的に表現する。このような定性的知識を用いたESの診断性能は、定量的な観測値を定性値に変換する写像関数に依存し、写像関数の最適化が重要となる。筆者らは既にシミュレーテッドアニーリング(SA)を用いて定性的知識と事例(定量値と本来出力されるべき診断結果の組)から写像関数を最適化する手法を提案した^[2]。本稿では、疎通悪化診断に用いるESを対象にして、提案の手法で写像関数を最適化した場合の効果を測定し、同手法の有効性を実証する。

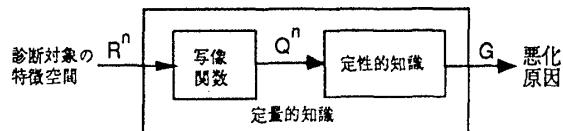
2. シミュレーテッドアニーリング(SA)を用いた写像関数最適化手法^[2]

n 個の定量値を写像関数で定量化した長さ n の記号列に定性的知識を適用することで診断結果を得るシステム(図1)を考える。写像関数は、図2に示すよう、事前に設定した境界値との大小比較によって変数毎に独立に定性値に変換するものとする。この時、写像関数は境界値の順序集合 θ^i で定義できる。

θ^i の善し悪しを評価するための評価関数 $E(\theta^i)$ は、正しい診断ができない確率と、悪化原因の分布状況を定性値で符号化する時の情報量に重みを乗じたものの和であり、与えられた定性的知識と事例から計算される。診断が正しくなるほど、また、悪化原因の分布を定性値でより簡潔に表現できるほど、 $E(\theta^i)$ の値は小さくなる。 $E(\theta^i)$ の値を最小にする θ^i が最適な写像関数である。

θ^i を求める際、適当な粒度の格子を考え、境界値

は格子点上にのみ存在するとした場合でも、解空間は非常に大きい。 θ^i を解候補、 $E(\theta^i)$ をエネルギー関数、隣接格子点を近傍としてSAを適用し、与えられた定性知識と事例から最適な写像関数を得ることができる。



Rⁿ: 実数(定量値)の組 Qⁿ: 記号(定性値)の組 G: 記号

図1: 定性的知識を利用する診断システムの構成

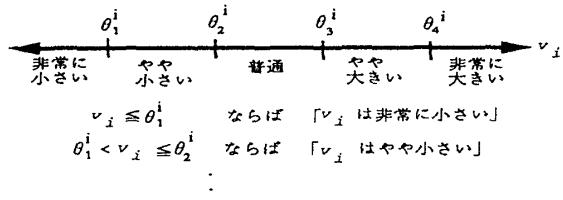


図2: 変数 v_i を 5 つの定性値に変換する例

3. 評価実験

3.1 使用した定性的知識と事例

国際電話網の疎通診断を行うES^[1]の知識ベースから、表1に示す3種類の定性的知識(KB1, KB2, KB3)を取り出し、評価に用いた。各定性的知識は、与えられた定性値の組み合わせから、悪化原因の候補(複数)を出力する。

また、実際に起こった疎通悪化に対し、診断に用いた定量値とその時の本当の悪化原因の組を事例とし、216件用意した。

3.2 実験内容

実験1 専門家が定めた写像関数 M1, M2, M3 にそれぞれ KB1, KB2, KB3 を適用し、診断性能として、出力に正しい診断結果を含む割合(Inc)と出力する候補の平均値(Comb)を測定した。

実験2 M1, M2, M3 を使用することを前提に、先の全事例を用い、候補として本来出力されるべきであった悪化原因が必ず出力されるよう修正を加えた定性的知識 KB1', KB2', KB3' を作成した。

"Evaluation of Optimization method for Mapping Function from Qualitatives to Quantitatives with Simulated Annealing", Kazunori MATSUMOTO, Kazuo HASHIMOTO, and Sadao OBANA: Kokusai Denshin Denwa Co. Ltd.

専門家が定めた写像関数と修正した知識の組み合せで診断性能を測定した。

実験3 SA を用い、全事例と KB1, KB2, KB3 から最適化した写像関数 M1', M2', M3' を求め、それぞれの診断性能を測定した。

3.3 実験結果

実験1,2,3の結果を表2に示す。実験1と2を比較すると、定性的知識の修正によって Inc が向上しても Comb は悪くなっている。それに対し、実験1と3とを比較すると、写像関数を最適化したことによって、どの定性的知識を適用しても、Inc も Comb も良くなっている。

表 1: 使用した定性的知識

	変数の内容	診断への影響
KB1	複数経路に分散したトラヒックを合算したもの	大局的な判断であり、出力の曖昧性は少ない
KB2	単一経路を通る通信相手へのトラヒック	経路毎の局所的な判断であり、出力の曖昧性は多い
KB3	通信相手以外のトラヒックを含む単一経路のトラヒック	経路毎の局所的な判断であり、出力の曖昧性は中程度

表 2: 診断性能

		Inc(%)	Comb	$E(\theta)$
実験 1	M1+KB1	84.7	3.82	1.56×10^{-1}
	M2+KB2	100.0	10.04	1.33×10^{-3}
	M3+KB3	99.0	5.04	1.14×10^{-2}
実験 2	M1+KB1'	100.0	4.65	2.88×10^{-3}
	M2+KB2'	100.0	10.04	1.33×10^{-3}
	M3+KB3'	100.0	5.20	1.68×10^{-3}
実験 3	M1'+KB1	100.0	3.78	9.54×10^{-4}
	M2'+KB2	100.0	9.96	5.96×10^{-4}
	M3'+KB3	100.0	3.62	8.78×10^{-4}

注: KB2 と KB2' は同一

4. 考察

(1) 提案の写像関数最適化手法の場合、正しい診断ができない確率が小さい写像関数の方がより最適なものとなる。実験1に比べ、実験3の Inc の方が高くなっていることから、同手法によって正しい診断ができない確率がより小さくなる写像関数を見つけることができたと言える。

(2) 同手法は、情報量に基づいた最適化尺度を持っており、悪化原因の分布が簡潔に表現できる写像関数ほど、より最適な写像関数となる。図3は、

この情報量に基づいた最適化尺度が有効に働き、定性知識の出力の曖昧性が減る例である。同図の(A)と(B)を比べると、(B)の方がより最適な写像関数であり、出力の曖昧性も小さい。実験2と3を比較すると、Inc が等しくても、実験3の方が Comb と $E(\theta)$ の値が小さい。これは、情報量に基づいた最適化尺度によって出力の曖昧性が減ったためと考えられる。

(3) 提案の写像関数最適化手法は、その適用対象となる定性的知識の性質や特徴空間上での事例の分布状況を仮定していないため、適用対象に対する強い制約はない。また、今回の実験結果は、同手法が表1に示したように扱う情報や診断の性格も異なっている3種類の定性的知識に対して有効であったことを示している。よって、手法自体の性質と実験結果から、同手法有効範囲は広いと考えられる。

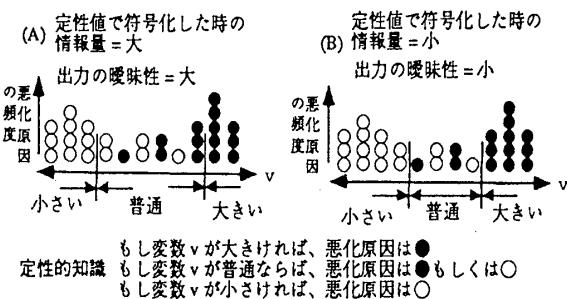


図 3: 定性知識の出力の曖昧性が減る例

5. おわりに

一般的に、定性的に表現したルールの信頼性は高くとも、それを定量的に表現すると信頼性は低下しやすい。これは、定性値に対応する実世界の定量値の範囲を適切に定めることが難しいからである。既に提案の SA を用いた写像関数の最適化手法によって、診断性能を向上できることが今回の評価実験によって確認できた。最後に、日頃御指導頂く KDD 研究所浦野義頼所長、眞家健次次長に感謝します。

参考文献

- [1] 松本等: 国際電話網の悪化検出/原因推定のための網運用支援エキスパートシステム, 信学技報 SSE-93-129, pp.49-54 (1994)
- [2] 松本等: シミュレーテッドアニーリングを用いた定量的知識から定性的知識への写像関数の最適化, 情処全大 48回前期, 5R-10, 3, pp.215-216 (1994)