室蘭工業大学
学術資源アーカイブ
Muroran Institute of Technology Academic Resources Archive

設計者の要求に基づく非劣解分析支援システムの提案

メタデータ	言語: Japanese 出版者: 計測自動制御学会 公開日: 2016-02-17 キーワード (Ja): 非劣解分析, 相関分析, オンデマンド キーワード (En): 作成者: 中野, 翔, 渡邉, 真也, 千葉, 一永, 金崎, 雅博, 棟朝, 雅晴 メールアドレス: 所属:
URL	http://hdl.handle.net/10258/3863

室蘭工業大学
学術資源アーカイブ
Muroran Institute of Technology Academic Resources Archive

設計者の要求に基づく非劣解分析支援システムの提案

著者	中野　翔，渡邉　真也，千葉　一永，金崎　雅博，棟朝　雅晴
雑誌名	計測自動制御学会システム・情報部門学術講演会講演論文集
巻	2015
ページ	631-635
発行年	2015-11-18
URL	http://hdl.handle.net/10258/3863

設計者の要求に基づく非劣解分析支援システムの提案*

○中野 翔 (室蘭工業大学　情報電子工学系学科)
渡邉 真也 (室蘭工業大学大学院 しくみ情報系領域)
千葉 一永 (電気通信大学大学院 情報理工学研究科)
金崎 雅博 (首都大学システムデザイン研究科)
棟朝 雅晴 (北海道大学情報基盤センター)

概要 多目的最適化によって得られた非劣解集合に対して設計者側の要求に基づいた傾向分析を実現する新たな分析支援システムの提案を行う．本システムは，従来までの相関ルールに基づく分析アプローチを改良したものであり，設計者側からの要求に即したルール抽出に特化することで，設計者の知りたい傾向をピンポイントで抽出することが可能となっている．

キーワード: 非劣解分析，相関分析，オンデマンド

1　はじめに

進化アルゴリズムを多目的最適化問題へ適用した進化型多目的最適化 (Evolutionary Multi-criteiron Optimization:EMO) では，MOEA/D[1]，NSGA-Ⅲ[2] といった高性能アルゴリズムの登場および計算機環境の飛躍的向上に伴い様々な分野への応用，4 目的を超える多数目的問題への適用など新たな盛り上がりを見せている [3, 4, 5]．

一方，このような様々な応用事例および多数目的問題への EMO の結果である非劣解集合からどのように問題の特性や傾向を読み取れば良いのかという非劣解分析が大きな課題として浮かび上がってきた．EMO により得られた非劣解集合は,「他の解と比較して一意に劣っていない」という特徴を有する非常に有益な情報のエッセンスとして捉えることができる．しかしながら，多変数，多目的かつその解候補の数も多数存在するためその解析が困難である．

この問題について我々は，相関ルールを用いた分析アプローチの提案を行い，多粒度な傾向の分析を可能にした [6]．提案手法は，これまでの非劣解集合の統計情報に基づく可視化，分析手法とは異なり，まず非劣解集合に対して論理分析を行う．その上で非劣解集合中に含まれる特性情報を適切抽出した上で，抽出した情報に焦点を絞り可視化，分析するというアプローチをとる．

一方，前述の先行研究では，全体の大きなトレンドを提示することは出来るものの，設計者が求める特定の目的に限定的な情報を必ずしも提示できるわけではないという問題を含んでいる．そこで，本研究では，この問題を解消することを目的に，先行研究を設計者の要求に基づく分析結果を出来る限り提示するシステムへと改良を行った．具体的には，相関ルールから設計者の可視化したい目的関数に対して支持度基準を緩めることで，先行研究では提示が不可能であった非劣解集合の大部分に関わる広域的な分析情報からごく限られた部分にのみ関わる特殊性の高い分析情報までの提供を実現している．

2　論理分析

提案する非劣解分析手法では，非劣解集合の持つ特性や傾向を論理分析を用いて抜き出し，その結果に基づいて情報の整理，再構造化を行う．そのため，論理分析は本提案手法の非常に重要なテクニックである．ここでは，論理分析における最も主要な概念である相関ルール，およびその代表的な導出アルゴリズムである Apriori アルゴリズムについて簡単に説明する．

論理分析は，離散値のみを扱い相関ルールを抽出する．そのため，対象が連続値であった場合には，まず離散化という処理を施した上で，後述する Apriori アルゴリズムを用いて相関ルールの抽出を行う必要がある．

2.1　相関ルール

相関ルールとは，ある事象が発生すると別の事象が発生しやすいという共起性のことであり，$A \rightarrow B$ という相関ルールは，A という事象が起こると B という事象が起こりやすいことを意味する．

相関ルールの重要性を測る指標として支持度と確信度が使われる．支持度とは，ルールが全トランザクションデータでどの程度出現するかを表す割合であり，式 (1) で表される．トランザクションとは 1 つのデータのことで，各トランザクションに含まれている項目をアイテムと呼ぶ．

$$\text{支持度}\ (A \rightarrow B) = \frac{A\ \text{と}\ B\ \text{がともに出現するトランザクション件数}}{\text{全トランザクション件数}} \tag{1}$$

支持度の高いルールは出現頻度が高い重要なルールであると言える．一方確信度とは，条件部が起こったときに結論部が起こる割合であり，式 (2) で表せる．

$$\text{確信度}\ (A \rightarrow B) = \frac{A\ \text{と}\ B\ \text{がともに出現するトランザクション件数}}{A\ \text{を含むトランザクション件数}} \tag{2}$$

確信度の高いルールは，条件部と結論部の結びつきが強く信頼できるルールと言える．また支持度が高いルールはその正否に関わらず，全体に当てはまるルールと言える．その条件部がほとんど出現しない場合は全体であまり出現するルールではないので，確信度だけが高くても意味がない．すなわち，ルールの重要性を判断する上で確信度が高いことは必須であるが，支持度もある一定以上の高さが必要であるといえる．

本手法では単にアイテムの組み合わせからルールを生成するので，確信度と支持度にある一定の閾値を設けなければ，極めて多数の無意味なルールが生成され

*本研究は SSI2015 で発表したものである．

SY0010/15/0000-0631 © 2015 SICE

てしまう．そのため相関ルールの抽出では支持度と確信度にある閾値を設け，その値以上の支持度と確信度を有する相関ルールのみを抽出するのが一般的である．それぞれの閾値を最小支持度，最小確信度といい，最小支持度以上の支持度を有するアイテム集合を多頻度アイテム集合と呼ぶ．

本研究では，論理分析ソフトウェアである Weka を使用している[1] [7]．

3 相関分析に基づく非劣解分析手法

ここでは，非劣解集合に対する分析支援として，本研究の元となっている相関分析を利用した先行研究 [6] について述べる．

情報視覚化の設計方針について，Shneiderman は “Over first, zoom and filter, then details on demand.” という 3 つのレベルで表現しており [8]，全体の概略から詳細なマクロ部分の分析までをユーザーの要求に従って提示できるような可視化であるべきとの指針を示している．

先行研究では，上記の指針要求を考慮したアプローチとして，論理分析に基づき非劣解集合を複数のグループに分割し，共通のルールに該当する非劣解集合のみを可視化，分析する多粒度に対応した分析支援システムを提案した．その特徴は，特徴抽出に論理分析を利用している点，得られた相関ルール同士をグループ化することにより多粒度な分析を実現している点，ユーザの興味の対象である目的関数に絞った分析を実現している点である．相関ルールを扱うことで，従来の研究では読み取ることが困難であった解の関係性および特性を自明のものとして扱うことができる．また，共通した性質を保持する相関ルール同士をグループ化することで，ミクロからマクロのどの粒度においても分析対象となる解集合の特徴および性質をある程度既知として分析することが可能としている．

3.1 非劣解集合の可視化

目的関数の数が 4 目的を超える場合，非劣解集合を直接可視化することができないため何らかの低次元化が必要となる．また，小数目的の場合においても設計変数次元を含めた解の次元で考えた場合，3 次元を超える超高次元となるため，解の持つ情報全体を把握するためにはやはり何らかの低次元化が必要となる．

千葉らは論理分析を用いて，多目的最適化の結果である非劣解集合から相関ルールを導出することで，可視化を行う方法を実現している

3.2 非劣解集合の離散化

連続値の離散化とは，対象の数値が取りうる最大値と最小値の区間を指定数で分割し，記号で置き換える処理を指す．離散化は，データに内在する情報を欠損させる危険性があるだけでなく，得られる相関ルールの数および質にも大きく影響を及ぼす非常に重要な役割を持っている．

最も単純な離散化手法としては，最大値と最小値間を一定間隔で分割を行う等間隔区間と一定のデータの個数ごとに分割を行う等頻度区間があるが，これらの手法ではデータの傾向を無視した区間決定を行うため，元のデータの性質が大きく損なわれる危険性がある．そこで，本研究では，次に述べるデータの分布に基づく方法を採用した．

データ分布に基づく離散化

データの離散化は，変数ごとに行われるため原則，1 次元空間が対象となる．データ分布に基づく離散化の手順を以下に示す．

Step1: 対象データを昇順にソーティングし，隣り合うデータ間の距離を測定．

Step2: データ間の距離の大きい区間順に指定分割数の数だけ選び，それらを離散化における区分として使用し，離散化を実現．

上記の手法により，単純な等間隔，等頻度に比べ，データの偏り具合を反映させた離散化を実現できるとしている．

3.3 相関ルールのグループ化と階層化

先行研究では，求めた相関ルール同士の類似性に基づく統合を行うのではなく，全相関ルールに含まれる設計変数及び目的関数を抽出し，それらの組み合わせを条件部もしくは結論部に持つ相関ルールを探索し，グループを生成するという方法を用いた．この方法を利用することにより，統合によるグループの不明瞭化のリスクを回避できる上，組み合わせ間の包含関係から階層構造を容易に生成することができる．具体的な手順を下記に記す．

Step1: 非劣解集合全体より結論部に目的関数を含む相関ルールを抽出．

Step2: 個々の相関ルールに含まれる設計変数及び目的関数の離散値を抽出し，重複を持たないすべての組み合わせを算出．

Step3: Step2 で求めた設計変数及び目的関数の離散値の組み合わせを条件部もしくは結論部に持つ相関ルールを探索し，グループを生成．

Step4: Step2 で求めた設計変数及び目的関数の離散値の組み合わせ間の包含関係から Step3 で求めたグループを階層構造化．

上記手順の概念図を Fig. 1 に示す．

では，Step1 により得られた 5 個の相関ルールに含まれる設計変数 A2, B4，C2 と目的関数として A1,D1 の 5 つが抽出されたと仮定している[2]．まず，Step2 において，A1, A2, B4, C2, D1 の全ての組み合わせについて算出し，Step3 ではそれらの組み合わせに該当する相関ルールを検出しグループ生成を行っている．Step4 では，A, B, C, D の組み合わせにおける包含関係に基づきグループの階層化を行い，階層構造を持つグループ化を実現している．先行研究では，すべての離散値の組み合わせを階層構造化している．

4 先行研究のオンデマンド対応への改良

一般に支持度を下げ，相関ルールを導出した場合には，全体的な傾向から外れた微細な傾向も導出されるため，ルール数が膨大になり，必要な情報に着目することが難しい．先行研究ではこのことを下げるため，一定以上の支持度を閾値をして設けており，導出ルールを一定数に留めている．このことにより，得られるルール数が限定され，抽出された傾向の分析は容易になる

[1] DataMiningSoftwareWeka

[2] A1 は目的関数 A を離散化した際の 1 番目のカテゴリを意味し，B4 は変数 B の 4 番目の区間を意味する．

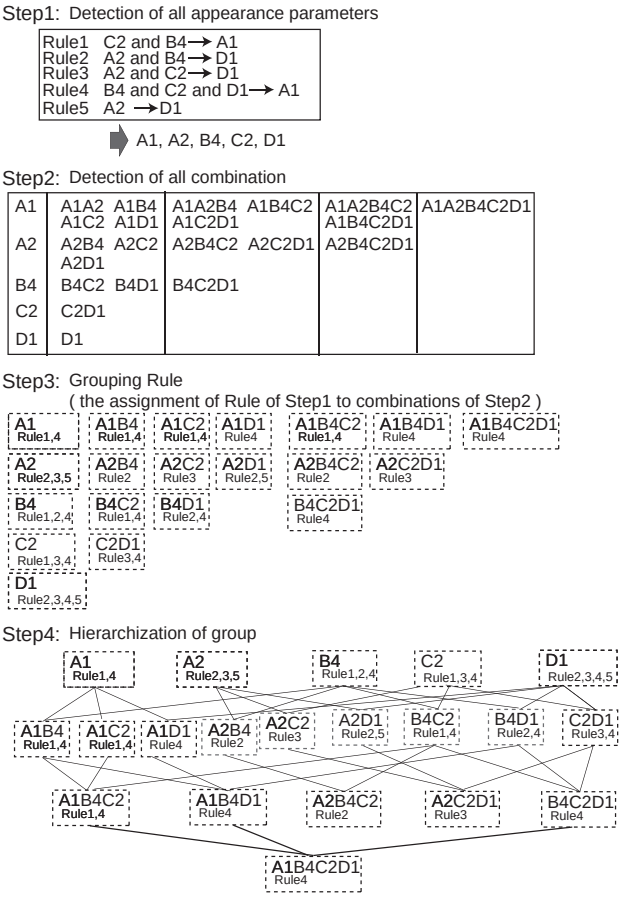

Fig. 1: The concept of grouping rules.

一方，微細な傾向は無視され，設計者の望む情報の取りこぼしにつながっていた．そこで本研究では，膨大なルールの中から要求に沿ったルールのみを厳選するフィルタリング操作により，得られたルール集合から要求に沿った情報だけを抜き出し，分析対象とすることで，過不足なく情報を導出することを試みた．また出力も，目的関数のみに特化することで設計者に必要な情報のみを提示している．

以下，提案するシステムの詳細について述べる．

4.1 システムの手順

提案システムの手順を Fig. 2 に示す．本システムでは，設計者の要求目的関数の傾向を導出するため，支持度を下げて相関ルールの導出を行う．まず，設計者の設定した分割数に基づいて離散化を行い，設計者の望む目的関数とその目的関数の離散化後の区間を設定する．Weka によって導出された相関ルールは支持度が低いため膨大となる．また，要求外の目的関数や設計変数においては，考慮せず出力するため，要求に基づきフィルタリングを行い，その後グループ化・階層化を行う．

以下，提案システムにおける離散化と相関ルールのグループ化および階層化・フィルタリングについての特徴を述べる．

4.2 離散化

3.2 説で述べたように，離散化はデータ傾向を読み取る上で重要な要素になる．本システムでは，設計者に離散化によるデータを感覚的に伝え，区間設定を行うため，等間隔による離散化を採用している．

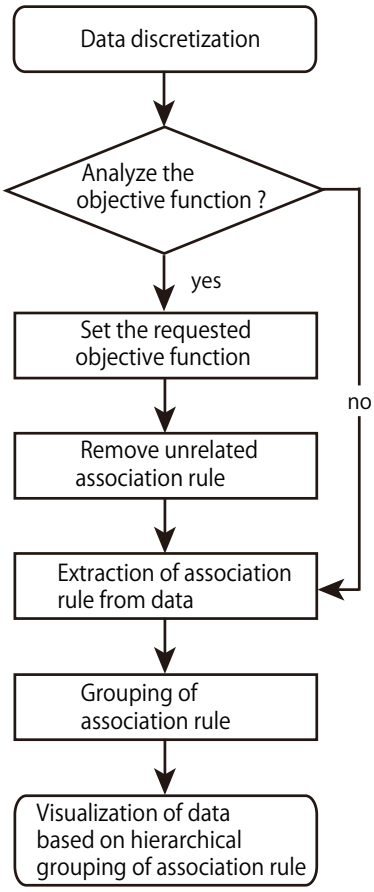

Fig. 2: Flow of Proposed System.

等間隔離散化

データの離散化は，等間隔に行われるため原則，1 次元空間が対象となる．データ分布に基づく離散化の手順を以下に示す．

Step1: 対象データの最大値，最小値，分割数から区間の間隔を求める．

Step2: Step1 の区間距離から離散化を実現．

4.3 相関ルールのグループ化と階層化

グループ化には，先行研究と同じように，相関ルールに含まれる設計変数・目的関数の離散値の組み合わせを利用する．離散値には目的関数・設計変数の種類とその区間の情報が含まれており，本システムでは組み合わせ内に指定された種類・区間が含まれる組み合わせのみを厳選し利用する．

Step0: 設計者の要求として，分割数，興味のある目的関数・区間を設定．

Step1: 非劣解集合全体より結論部に目的関数を含む相関ルールを抽出．

Step2: 導出された相関ルールから要求の含まれないルールを取り除く．

Step3: 個々の相関ルールに含まれる設計変数及び目的関数の離散値を抽出し，重複を持たず，設計者の要求を満たす組み合わせを算出．

Step4: 求めた離散値の組み合わせを条件部もしくは結論部に持つ相関ルールを探索し，グループを生成．

Step5: Step3 で求めた設計変数及び目的関数の離散値の組み合わせ間の包含関係からグループを階層

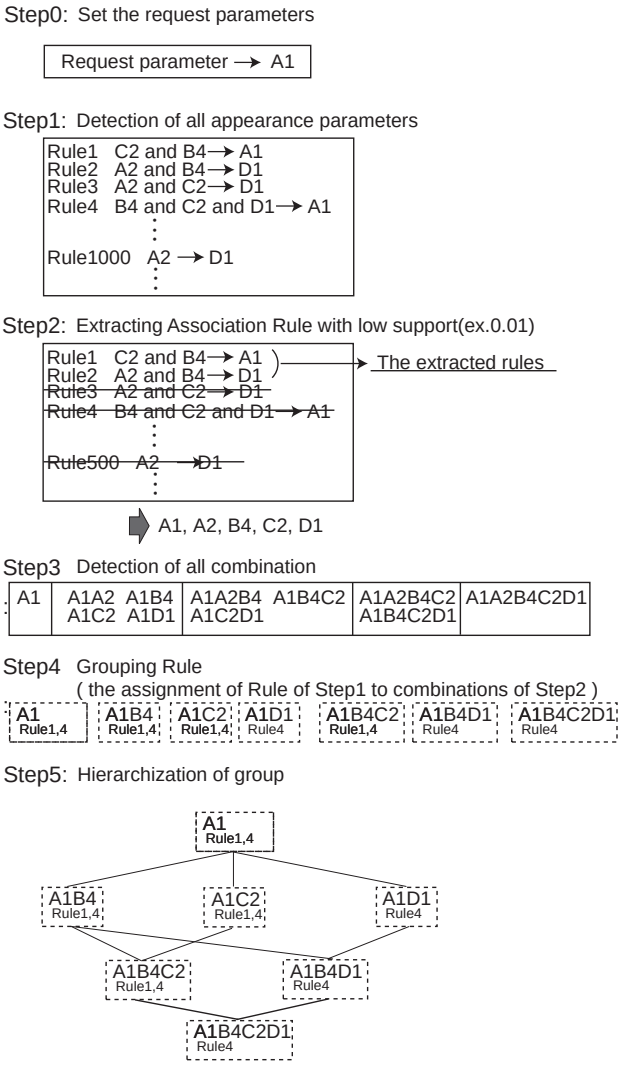

Fig. 3: The concept of proposed approach.

構造化.

上記の手順に関する概念図を Fig. 3 に示す．Fig. 3 では，4.3 節と同じく，設計変数 A2，B4，C2 と目的関数として A1，D1 の 5 つが抽出されたと仮定している．本システムでは，Step0 で設定された要求により得られた相関ルールから，ユーザに指定された目的関数である A1 についてのルールのみを抽出している．

本研究では，Step2 で要求目的関数・区間が含まれる相関ルールのみを抽出することで，あらかじめ無駄な情報を省くことで，より厳選された情報を提示するとともに，支持度を下げたミクロな結果を導出することも可能にしている．また，目的関数の区間については，分割数が増えると導出される相関ルールも増加する傾向にあるため，複数選択も可能となっている．

5 数値実験

数値実験として，進化型多目的最適化により得られた，1000 個の非劣解集合への提案手法の適用を試みた．対象問題は 3 目的 9 変数のハイブリッドロケットエンジン設計問題 [9] であり，ここでは，下部熱圏でのダウンレンジの最小値付近の傾向について，設計者側からの要求があったと仮定し，従来手法による分析と今回の拡張手法との比較を行った．

今回のハイブリッドロケットエンジン設計問題は Table. 1 のような目的関数・設計変数を持つ問題である．それぞれの変数の分割数を 10，最小支持度を 0.05，設計者の要求が下部熱圏でのダウンレンジが区間 [0,1,2] をとる場合を仮定し実験を行う．また，下部熱圏でのダウンレンジ (目的 0) を以降便宜上 of0 と呼ぶ．

先行研究で得られた結果を Fig. 4 に示す．先行研究の結果は，階層木構造では 698 ノードのグループが得られているものの，グラフの密度が濃く，視覚的に目的と変数の関係を理解することが出来ない．また，導出されているグループの中には，of0=[0,1,2] がまったく含まれていないグループもあり，軽微な傾向を読み取ることは不可能である．

提案手法での結果を Fig. 5 ，Fig. 6 に示す．Fig. 5 は要求目的関数・区間に基いた相関ルールのフィルタリングを行った結果である．この結果は，階層木構造では 244 ノードのグループを持ち，先行研究に比べるとノード数が約半分となっている．全体のグラフの見易さは先行研究に比べると，向上しており，設計者の望む目的関数に関する全体の傾向を読み取りやすくなっている．結果は煩雑になることもあるが，設計者の要求に関係しない部分は出力されていない結果になることから，可視化の結果すべてが設計者の求める情報であり，過不足ない結果であるといえる．

Fig. 6 はグラフ内の 1 番目の区間について注目し，of0=[0] の区間をとるときの組み合わせ以外を排除したものである．この結果は要求区間を，さらに細かく分けることで可視化を行った際の結果をみることが出来る．実験結果として，階層木構造でのノード数は 72 ノードとなっており，先行研究の結果の約 10 分の 1 まで，ノード数を省略しており，目的関数と設計変数の関係が視覚的にわかるようになっている．また，前述の通り，Fig. 6 は Fig. 5 の一部を拡大したものであるが，注目する区間に関係する部分のみを出力することから，よりミクロな視点から分析を行うことが出来る．

この結果から，支持度を下げるだけでは，設計者の求める情報を得ることは出来ず，要求に応じて適切な情報を選択することで，傾向をピンポイントで提示することが出来ることがわかる．

6 おわりに

本稿では，相関ルールに基づいた非劣解集合分析のオンデマンド手法への拡張の提案を行った．

具体的には，支持度を従来手法よりも大幅に下げ，より微細なルール抽出を行う一方，要求に沿った情報だけをフィルタリングすることで，設計者の求める情報をピンポイントで抽出する手法である．数値実験を通して，従来手法では，膨大なルール数から抽出できなかった要求された傾向のみの分析結果を導出することが出来るという点で，有用性を示す結果であったといえる．

今後は下記に示す課題について取り組む予定である．

- 相関ルール分析に基づくため，結果はルールの量，及び質に左右される．
- 新たな離散化の導入・開発．

参考文献

1) Q. Zhang and H. Li. Moea/d: A multi-objective evolutionary algorithm based on decomposition. *IEEE*

Table 1: Hybrid Rocket Engine design problem

Objective function	Name of the function
of0	downrange in the lower thermosphere(Maximization)
of1	duration time in the lower thermosphere(Maximization)
of2	initial gross weight of launch vehicle(Minimization)
Design variable	**Name of the design variable**
dv0	initial mass flow of oxidizer
dv1	fuel length
dv2	initial radius of port
dv3	total combustion time
dv4	1st combustion time
dv5	extinction time
dv6	initial pressure in combustion chamber
dv7	aperture ratio of nozzle
dv8	elevation at launch time

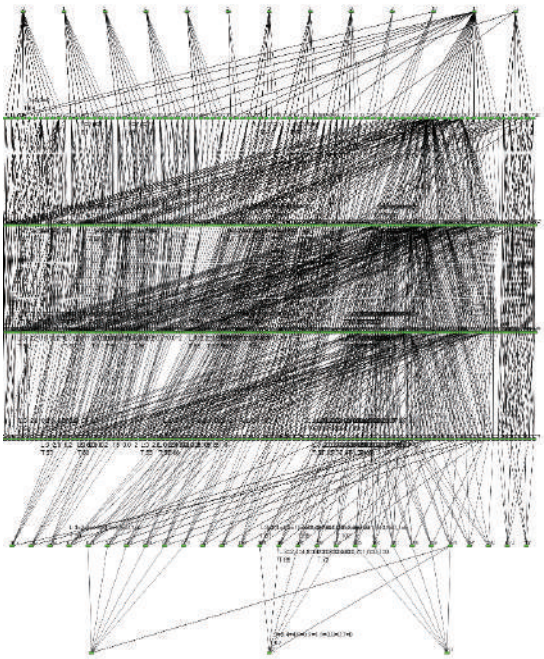

Fig. 4: The result of previous approach.

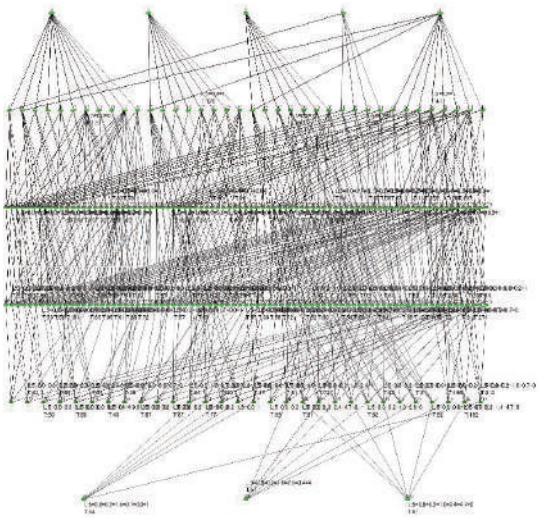

Fig. 5: The result of the proposed approach (of0 = [0,1,2]).

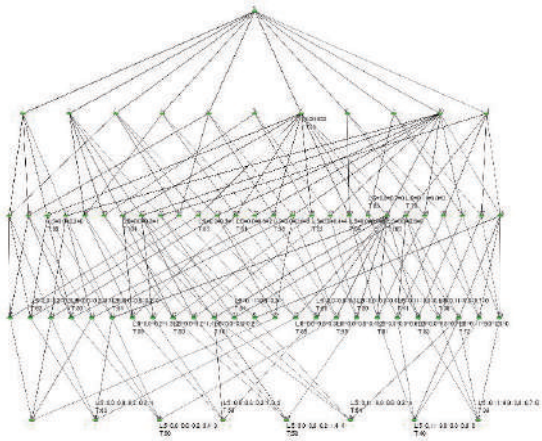

Fig. 6: The result of the proposed approach (of0 = [0]).

Trans on Evolutionary Computation, pp. 712–731, 2007.

2) K. Deb and H. Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problemswith box constraints. *IEEE Transactions on Evolutionary Computation, Vol. 18, No. 4*, pp. 577–601, 2014.

3) K.Deb. *Multi-Objective Optimization using Evolutionary Algorithms.* Chichester, UK:Wiley, 2001.

4) C. A. Coello Coello and G. B. Lamont. *Applications of Multi-Objective Evolutionary Algorithms.* World Scientific, Singapore, 2004.

5) K.Deb. *Innovization: Innovative Solution Principles Using Multiobjective Optimization.* Springer-Verlag New York, Inc., 2012.

6) 渡邉真也, 千葉祐大, 金崎雅博. 相関ルール分析に基づく非劣解分析支援システムの提案:ハイブリッドロケットエンジンへの適用を通して. 人工知能学会論文誌, Vol. 29, No. 2, pp. 219–223, 2014.

7) Ian H. Witten and Eibe Frank. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems).* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.

8) B. Shneiderman. *Designing the User Interface (3rd Ed).* Addison-Wesley, 1998.

9) Shoma Ito Kazuhisa Chiba, Hideyuki Yoda and Masahiro Kanaza. Visualization of design-space constitution for single-stage hybrid rocket with rigid body in view of extinction-reignition. *submitted to 2015 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2015.