

# 多目的遺伝的アルゴリズムによる制限食を考慮した自動献立作成システムの開発と高速化

1915077 水上和秀

情報基盤工学講座

指導教員 António Oliveira Nzinga René

## 要約

本研究では、自動献立作成システムを、生活習慣病を患った人やアレルギーを持っている人でも利用できるように改善することを目的とする。また、献立作成の解を得るために膨大な時間がかかることを考慮し、並列分散処理をはじめとしたあらゆる処理を施しプログラムの高速化を図る。

キーワード：健康、献立作成、多目的最適化、並列分散

## 1 はじめに

戦後の食生活が欧米化したことに伴って、ファストフードといった、過剰にエネルギーを摂取してしまうような食生活が大きく広まったことから、現在、生活習慣病を患う人々が増加している。その要因には、偏った栄養による食生活や運動不足、飲酒、禁煙などが挙げられる。また、生活習慣病を患った場合、食生活を見直すことによって改善することができる[1]。しかし栄養バランスの取れた献立を作成するには、メニューの組み合わせや栄養価の計算を考慮する必要がある。献立を考えることは面倒だと考える人は少なくない。そのため、栄養やメニュー組み合わせの観点から献立そのものをコンピュータによって最適化し、自動作成する研究が存在する。本研究ではメニューの組み合わせや並列処理の変更を施し、プログラムの処理の高速化を図る。

## 2 自動献立作成システムの概要

### 2.1 Web上のレシピを活用した献立作成

献立作成を行うにはレシピデータが必要である。そのためインターネット上からレシピデータを取得する必要がある。Web上にはcookpadやクラシル、ボブとアンジーなどの料理レシピサイトが多数存在する。これらのサイトには、料理名、料理のイメージ、料理をするために必要な材料、得られる栄養素などの情報が掲載されている[3]。ウェブサイトから収集するデータのイメージを図1に示す。

また、生成食品や加工食品、畜産品などの価格動向を前月や前月同年と比較して提供しているサイトが存在している[4]。



図1 データの収集と活用

先行研究においてはレシピサイトである「ボブとアンジー」においてからレシピ情報と食材を、食品価格動向を調査しているWebサイトから食材の価格をスクレイピングしデータベースに蓄積し、献立作成の最適化に活用した。レシピサイトと食品価格サイトからスプレイングによって得られるデータの流れを図1に示す[2]。また、本研究もWeb上にあるレシピサイトをスクレイピングし、出力するレシピとして参照する。

### 2.2 最適な献立の提示

献立作成は多目的最適化問題としてとらえることができる。多目的最適化とはある制約条件のもと、複数の目的関数を最大化、あるいは最小化する最適化法である。多目的最適化問題は、あらゆる分野において存在し、それらは全ての目的関数を最大化、あるいは最小化するような最適解が存在するとは言えないため、パレート最適という概念を導入する必要がある。

献立作成システムにおける多目的最適化の定式化を図2に示す。目的関数には調理時間、調理コストの最小化を設定し、制約条件には、1日に摂るべき各栄養素の量やカロリーを満たすような献立を作成するという制約や、1つの献立に対して主菜を1つ、副菜を3つとする制約などを設定する[2]。

また、多目的最適化を解く手法として遺伝的アルゴリズム (Genetic Algorithm: GA) がある。GAとは、近似解を探索するためのアルゴリズムである。GAは、解の候補であるデータを遺伝子で表した「個体」を複数体用意し、適応度関数によって計算された適応度の高い個体を優先して選択し、生命の進化過程である「交叉」、「突然変異」や淘汰などの操作を繰り返し行うことで最適解の探索をする。本研究では、GAを多目的最適化問題に拡張したNSGA2を用いる。NSGA2の特徴として非優越ソート、混雑度トーナメント選択がある。

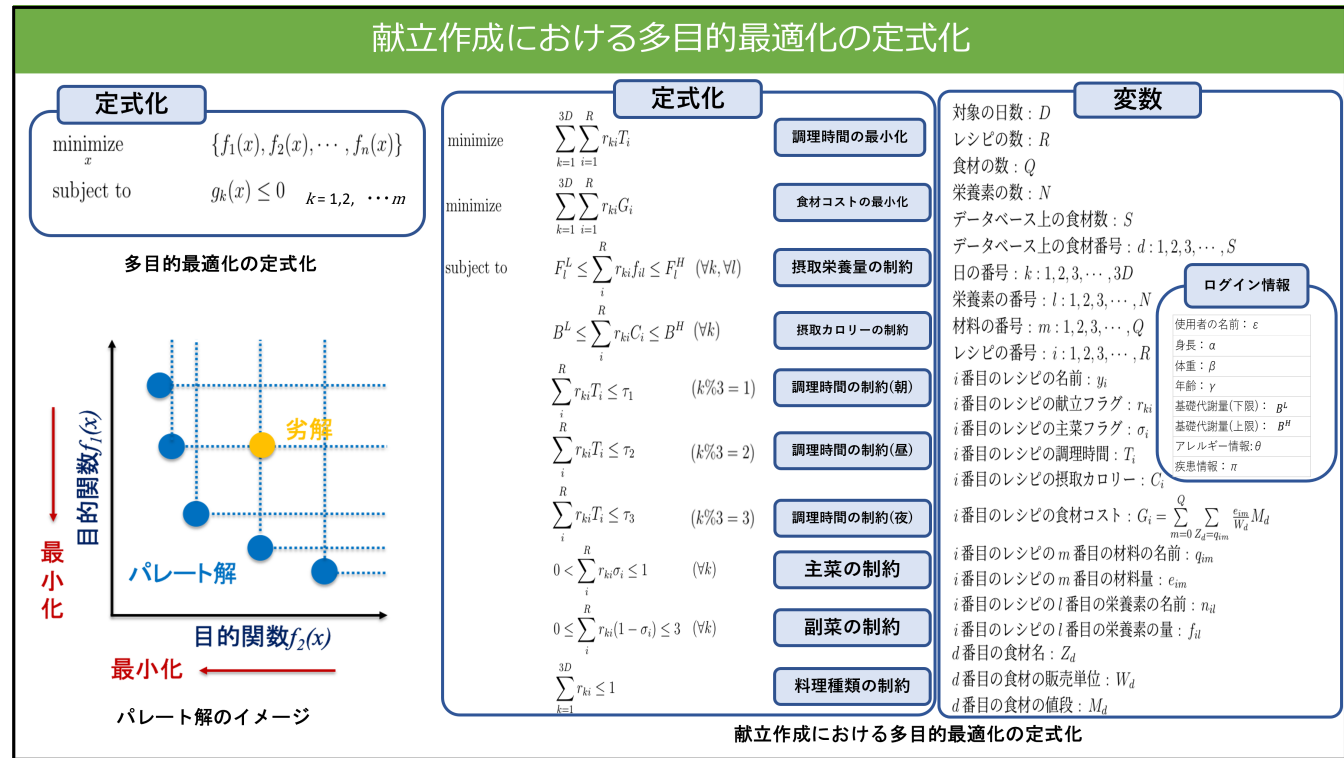


図2 多目的最適化の定式化

### 2.3 並列分散処理による解法

GAは、多目的最適化問題を解くときに複数の目的関数及び制約条件の値を繰り返し評価する必要があり、膨大な計算時間が必要となる問題がある。このため、並列処理により計算時間を短縮することは重要な課題となる。並列分散処理とは、複数台のコンピュータを用いて複数のCPUや、メモリを使うことで一つの計算処理を行い、性能や計算速度を向上を図ることである。本研究ではいくつかの数値実験を行い従来の並列分散モデルとの比較を通して提案するモデルの有効性の検証を行う。

## 3 制限食を考慮した遺伝的アルゴリズム

### 3.1 様々なレシピサイトに対応した献立の追加

レシピサイトによっては、掲載されているレシピの種類に偏りがある場合がある。先行研究でレシピサイトとして参考にしてきた「ボブとアンジー」には主食が掲載されていないため、出力結果に主食が出力されていないという問題が生じていた[2]。そこで本研究では「ボブとアンジー」に加え、「EatSmart」と「おいしい健康」の2つのサイトからスクレイピングを行い、多種多様なレシピを参照し、より実用的な献立作成を提案する。

「EatSmart」は株式会社EatSmartが運営しているレシピサイトであり、主に主食の栄養素を掲載している。また、「おいしい健康」は管理栄養士が考案したレシピが掲載されており、生活習慣病にあった制限食が掲載されている。「EatSmart」からは主食を、「おいしい健康」からは制限食をスクレイピングし献立をより実用化できることを目的とする。レシピデータのスクレイピング情報を図3に示す。



図3 他のレシピデータのスクレイピング情報

### 3.2 健康のための制限食の考慮

献立を作成するにあたって、人によってはアレルギーを含む食品や生活習慣病による制限食を考慮しなければならない。制限食とは、個人の健康状態、病気の状態に合わせてカロリーや塩分などを制限する食事のことである。病気の種類にあった制限食を摂ることで病気の症状を改善することができる。また、病気にかかっていなくても生活習慣病を予防することができる。

本研究は人によって配慮すべきアレルギーや疾患を考慮した献立を作成することを目的とする。対象とする生活習慣病は糖尿病、高血圧、脂質異常症、腎臓病とする。また、アレルギーの対象項目として「特定原材料等」に指定されている、えび、かに、小麦、そば、卵、乳、落花生 (ピーナッツ)、アーモンド、あわび、いか、いくら、オレンジ、カシューナッツ、キウイフルーツ、牛肉、くるみ、ごま、さけ、さば、大豆、鶏肉、バナナ、豚肉、まつたけ、もも、やまいも、りんご、ゼラチンの28品目とする[5]。

### 3.3 自動献立作成の実用化のための短時間化

先行研究においてはプログラムの高速化として並列分散処理を施しており、ライブラリはDaskを用いていた。DaskはPythonのライブラリの1つであるNumpyや、データ解析を支援するために、時系列データや数表を操作できるデータ構造とその演算を提供している、Numpyと同様にPythonのライブラリであるPandasを、Daskは簡単に並列・分散して処理を行うことが可能である。

しかし大量のデータを扱う場合、分散処理においてApache SparkがDaskよりも高速に処理できることが示

されている[6]。そこで本研究で使用する並列分散ソフトウェアとしてApache Sparkを用いることで処理の高速化を図る。Apache Sparkは大規模データを使用する場合に使われ、分散コンピューティングエンジンの中でもっとも高速なエンジンの1つである。

## 4 提案手法

本研究で提案する、制限食を考慮した自動献立作成システムの流れを図4に示す。まず最初に、献立作成の最適化に必要な、レシピデータと食材価格データを、Webサイトからスクレイピングし、データベースに蓄積する。次に、ユーザーが身体情報やアレルギー情報、患っている生活習慣病を入力する。その際にアレルギーや嫌いな食品が含まれるレシピをデータベースから削除する。そして蓄積されたレシピデータ、ユーザの身体情報を入力として、摂取栄養素やカロリーなどの制約条件のもと、調理時間、調理コストの最小化を目的関数に設定した最適化問題を、制約条件を考慮した遺伝的アルゴリズムによって解く。最後に、設定した日にち分献立をユーザに出力する。さらに、最適化の工程で複数のPCを利用し、並列処理を行うことでプログラム実行時間の速度向上を図る。



図4 自動献立作成の流れ

## 5 数値実験並びに考察

今回の実験ではレシピサイトとして「ボブとアンジー」「eatsmart」「おいしい健康」の3つのサイトを参考にし、入力情報として年齢を21歳、身長を167cm、体重を67kgとした。また、並列分散としてDaskをもちいた。用いた制約条件と出力までにかかった時間を図5に示す。自動献立作成システムによって出力した1週間分の献立は、設定した制約条件を満たしながら最適化されていることが分かる。しかし、PCを4台用いて並列分散処理を施しても2時間30分かかってしまうのでより効率的な手法を取り入れる必要がある。

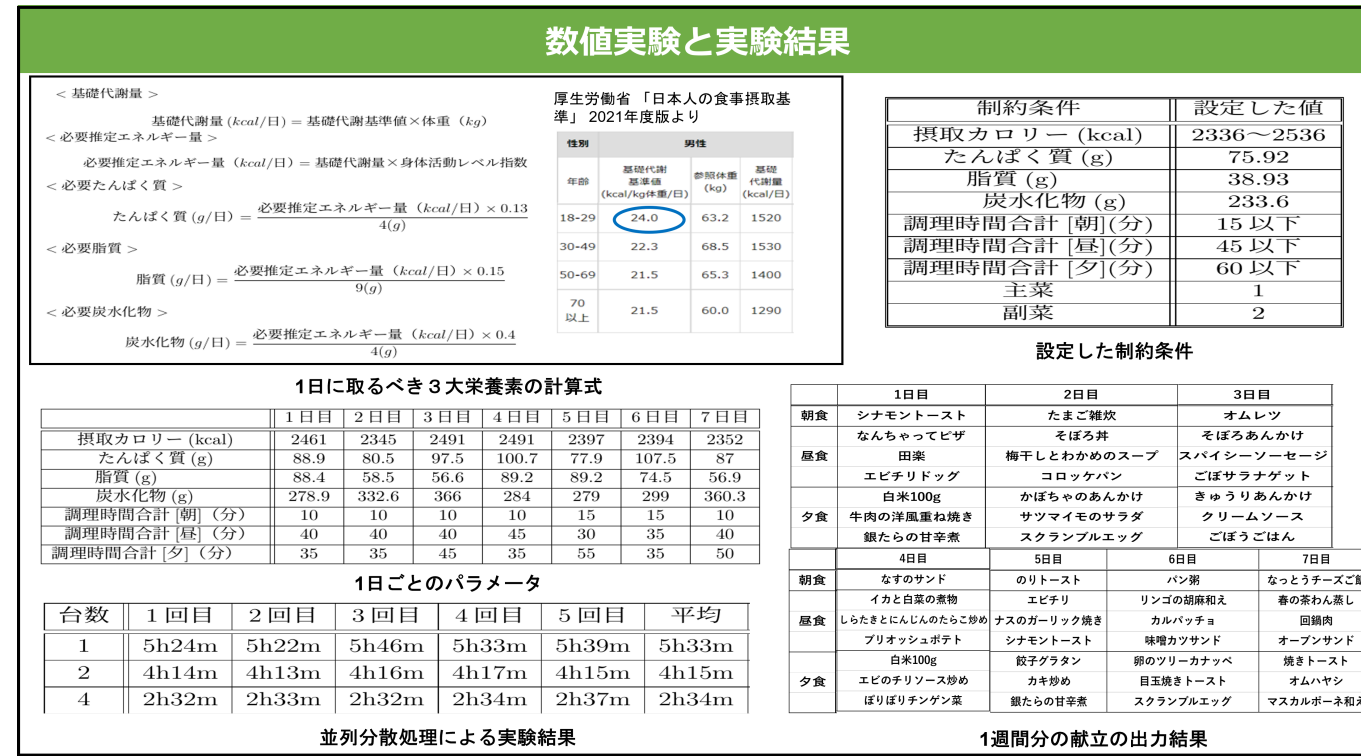


図5 実験結果

## 6 おわりに

今回は、複数のWebサイトのレシピ情報を参照し、レシピデータの多様化を図った。今後の方針として、制約条件を増やし、利用者によって疾患やアレルギーを考慮できるようにすることや、別の並列処理を施すことの処理の高速化などがあげられる。

## 参考文献

- [1] “生活習慣病(せいかつしゅうかんびょう)-e-ヘルスネット” <https://www.e-healthnet.mhlw.go.jp/informationdictionary/metabolic/ym-040.html>, 閲覧日 2022.11.2
- [2] 安藤祐斗, “Web 情報を活用した自動献立作成のための多目的遺伝的アルゴリズムによる並列分散処理” 富山県立大学学位論文, 2022.
- [3] “料理レシピ ボブとアンジー 管理栄養士監修の健康ヘルシーレシピ” <https://www.bob-an.com/>, 閲覧日 2022.11.2
- [4] “小売物価統計調査による価格推移” <https://jpmarket-conditions.com/>, 閲覧日 2022.11.2
- [5] “アレルギーを含む食品に関する表示, 消費者庁” <https://www.caa.go.jp/policies/policy/food-labeling/food-labeling-act/pdf/food/labeling/cms101/200716/12.pdf>, 閲覧日 2022.11.2
- [6] “Koalas (PySpark) がDask よりも高速な理由 - SQL クエリ最適化など” <https://www.databricks.com/jp/blog/2021/04/07/benchmark-koalas-pyspark-and-dask.html>, 閲覧日 2022.11.2