

大規模物流センターにおける配送スケジュールの基礎研究

旭川高専 ○高橋 麻希子 渡辺 美知子 古川 正志

要 旨

これまでに大規模物流センターをフローショップ問題として定式化し、LCO を適用する方法を開発してきた。1 台の AGV に 1 注文を割り当てる順序問題の解法として LCO は有効であったが、1 台の AGV (又は配送車) に複数の注文製品を最適に組み合わせる順序を作り出すことが課題として残された。本研究では、LCO を順序問題として適用するとともに、複数の注文製品を最適に 1 台の配送車に組み合わせる配送順序を強化学習から作成する配送計画法を提案する。

1. はじめに

交通網・情報網の発達により、生産システムではサプライ・チェーン・システムが導入され、また流通システムでは通信販売等の普及にともない、大規模な物流センターが各地に立地するようになってきている。このような物流センターでは、1 日 3~8 万件にも及ぶ多種類少品種の注文があり、こうした注文に対して納期にあわせる迅速な配送計画の作成が求められている。

これまでに、1 方向に移動起動をもつ物流センターを仮定し、1 台の AGV に 1 注文を割り当てる注文書の順序問題を生産計画問題におけるフローショップ問題として、局所クラスタリング組織化法(Local Clustering Organization, LCO)を適用した配送方法を開発してきた¹⁾。しかしながら、通常の物流センターでは、1 台の AGV に複数の注文が割り当てられるのが普通であり、複数の注文を 1 台の AGV に効率的に割り当てる配送計画を作成することが根本的な問題として残された。

本研究では、複数の注文を 1 台の AGV に割り当てる配送計画を作成するために、順序問題に LCO を適用するとともに、複数の注文を 1 台の AGV に割り当てる方法として注文製品の類似度を用いた、強化学習による方法を開発し報告する。以下では、2 章で物流配送センターでの問題を定式化し、3 章で LCO のアルゴリズムを説明し、4 章で(順最適割り当て)を求める強化学習法を述べ、5 章で提案するアルゴリズムとその一部の計算結果を示した上で、6 でまとめを行う。

2. 物流センター配送問題

通信販売物流センターでは、1 日に 3~8 万件の注文があり、納期にあわせて商品を梱包し、配送する大規模スケジューリングが必要とされている。しかしながら、商品には地域や季節により注文が集中するものとそうでないものがあり、従来のスケジューリングでは、センター内において一部が集荷に集中してしまうボトルネックが生じることがある。このような問題を解決するには、物流センターに注文の多い商品を分散させて配置する方法がひとつである。しかしながら、複数の注文を 1 台の AGV に割り当てる配送計画を作成する根本的な解決策とはならない。

本研究では以下のような問題を取り扱う。

$$\text{minimize } F(S) = \max_{S_i} F(S_i) \quad (1)$$

$$\text{subject to } S = \{S_i : i = 1, 2, \dots, n\} \text{ where } S_i = \{S_{i1}, S_{i2}, \dots, S_{im}\}$$

ここで、 $F(S)$ は AGV の最大滞留時間、 $F(S_i)$ は AGV S_i の滞留時間、 $S = \{S_i\}$ は AGV S_i の順序 (m 個毎の注文の順序)、 S_{ij} は AGV S_i に割り当てられた m 個の注文書の組

である。

本研究では、AGV の最大滞留時間 $F(S)$ を最小にするような S を決定するために LCO と強化学習を採用する。

3. 局所クラスタリング組織化法(LCO)

3.1. LCO の概要

LCO はリカッチ型学習方程式に基づいて、古川等が開発した局所最適化の繰り返しにより、大域的最適化を実現する方法である²⁾。順序問題では、局所最適化は局所的クラスタリングの結果となるので、局所的クラスタリングをランダムに実現することで大域的な(準)最適順序を得る。

3.2. LCO のアルゴリズム

$J = \{J_1, J_2, \dots, J_n\}$ の集合が与えられるとき、 J から作られる順序集合を $G = \{G_1, G_2, \dots, G_N\}$ とする。また、 $G_k = \{g_{k1}, g_{k2}, \dots, g_{km}\}$ とする。ここで、 $N = n!$ である。 $G_k \subset G$ である一つの順序に対して、その評価 $F(G_k)$ が計算できるとする。 G_k の要素は左から順に円と位相同型なトポロジを設定する。このとき、LCO のアルゴリズムは以下に示される。

- (1) $t=0$ とする。
- (2) J からランダムな順序の並びを生成し、これを初期解として $G_k(t) = \{g_{k1}, g_{k2}, \dots, g_{km}\}$ とする。
- (3) $g_{kh} \in G_k(t)$ なる要素を $G_k(t)$ からランダムに選択する。
- (4) g_{kh} の近傍要素集合 $N_h(t)$ を生成する。
- (5) $N_h(t)$ の要素を評価 $F(G_k(N_h(t)))$ に基づいてクラスタリングする。
- (6) $t=T$ または与えた条件が満たされれば終了する。さもなければ、 $t=t+1$ として(2)へ戻る。

上記アルゴリズムで、 $N_h(t)$ は g_{kh} から経路長 $d(t)$ 内にある $G_k(t)$ の要素集合である。

3.3. クラスタリングの手法

LCO のクラスタリングには、以下の 4 種類の方法が用いられる。なお、以下では、 $F(h+l, h+l+p)$ と $F(h+l+p, h+p)$ は前者が元の順序、後者が $h+l$ 番目と $h+l+p$ 番目の要素を交換した時の $G_k(t)$ の評価とする。

3.3.1 単純交換法(Simple Exchange Method, SEM) $G_k(t)$ からランダムに選択した h 番目の要素に関して、近傍 $N_h(t)$ 内の要素を c 番目の左右の要素順に、評価を計算してクラスタリングを実施する。アルゴリズムを以下に示す。

- 1) $l=0$ 及び $p=0$ とする。
- 2) $l=d(t)$ ならば、終了。そうでなければ p を $p+1$ とする。
- 3) $F(h+l, h+l+p) \geq F(h+l+p, h+p)$ ならば、解 $G_k(t)$ の $h+l$ 番目と $h+l+p$ 番目の要素を交換する。

- 4) $F(h-l-p, h-l) \geq F(h-l, h-l-p)$ ならば, 解 $G_k(t)$ の $h-l-p$ 番目と $h-l$ 番目の要素を交換する.
- 5) $l+p=d(t)$ ならば, l を $l+1$ とし, $p=0$ とする.
- 6) 2)へ戻る.

3.3.2 逆位交換法 (Inverse Exchange Method, IEM) $G_k(t)$ からランダムに選択した c 番目の要素に関して, c 番目の左右の要素順にそれまでの順序を逆にし, 評価を計算してクラスタリングを実施する方法であるこれは, $\{h, h+1, \dots, p-1, p\}$ の順序に要素が並ぶとき, これを $\{p, p-1, \dots, h+1, h\}$ のように並び替える方法である. 但し, 記号 $\{\cdot\}$ は, $G_k(t)$ の要素の順番のみを示すために用いた. このときの評価を $F(h \sim p)$ と $F(p \sim h)$ と表現する. IEM は以下のアルゴリズムとなる.

- 1) $p=0$ とする.
- 2) p を $p+1$ とする.
- 3) $F(h \sim h+p) \geq F(h+p \sim p)$ ならば, 解 $G_k(t)$ の h 番目から $h+p$ 番目の要素の順序を逆転する.
- 4) $F(h \sim p \sim h) \geq F(h \sim h-p)$ ならば, 解 $G_k(t)$ の $h-p$ 番目から h 番目の要素の順序を逆転する.
- 5) $p=d(t)$ ならば終了する. $p < d(t)$ ならば, 2)へもどる.

3.3.3 平滑法 (Smoothing Method, SM) R からランダムに選択した h 番目の要素に関して, N_h 内で最も左端 (N_h 内の1番目) の要素から順に比較を実施する方法である. アルゴリズムは以下となる.

- 1) $l=0$ 及び $p=0$ とする.
- 2) $l=2d(t)$ ならば, 終了. そうでなければ p を $p+1$ とする.
- 3) $F(h-d(t)+l, h-d(t)+l+p) \geq F(F(c-d(t)+l+p, h-d(t)+l))$ ならば, 解 $G_k(t)$ の $(h-d(t)+l)$ 番目と $(h-d(t)+l+p)$ 番目の要素を交換する.
- 4) $l+p=2d(t)$ ならば, l を $l+1$ とし, $p=0$ とする.
- 5) 2)へ戻る.

3.3.4 対称交換法 (Symmetrical Exchange Method, SYM) R からランダムに選択した c 番目の要素に関して, トポロジーの経路長が左右対称の位置にある要素をスケジュール評価で比較する方法である. アルゴリズムは以下となる.

- 1) $p=0$ とする.
- 2) p を $p+1$ とする.
- 3) $F(h-p, h) \geq F(h, h+p)$ ならば, 解 $G_k(t)$ の $h-p$ 番目と $h+p$ 番目の要素を交換する.
- 4) $F(h-p, h) \geq F(h, h-p)$ ならば, 解 $G_k(t)$ の $h-p$ 番目と h 番目の要素を交換する.
- 5) $p=d(t)$ ならば終了する. $p < d(t)$ ならば, 2)へもどる.

3.4 LCO の配送問題への適用

配送問題に LCO を適用するには, AGV の集合 $S = \{S_i\}$ を $J = \{J_i\}$ に対応させればよい. このとき, $S_i = \{S_{i1}, S_{i2}, \dots, S_{im}\}$ が注文の組合せとなるから, 全ての注文数は mn である.

4. 強化学習による AGV への注文の組合せ決定

4.1 Q 学習

LCO を配送問題に適応するためには, $S_i = \{S_{i1}, S_{i2}, \dots, S_{im}\}$ を決定する必要がある. この決定に強化学習に基づく規則を適用する. ここでは, 強化学習として Q 学習を採用する. Q 学習は, 過去の経験に基づく将来の行動期待値である. 今, 時刻 t の状態集合を $v(t)$, そのときの行動集合を $o(t)$, その時点状態を $v(t)$, 行動を $o(t)$ の Q 値を $Q(v(t), o(t))$ とする.

状態 $v(t)$ で行動を $o(t)$ を実行した結果, 時刻 $t+1$ で状態 $v(t)$ から状態 $v(t+1)$ へ移行したとき, $Q(v(t+1), o(t))$ は, 次式により更新される.

$$Q(v(t+1), o(t)) = (1-\alpha)Q(v(t), o(t)) + \alpha(R + \gamma \max_o Q(v(t+1), o(t))) \quad (2)$$

ここで, α は学習率, γ は減衰率, R は環境からの直接報酬, 関数 \max は行動集合 o の中から, その状態に対する最大の Q 値を得る関数である. このとき, 行動の選択は,

$$o(t) = \operatorname{argmax}_o (\exp(Q(v(t), o(t))) / \sum_o \exp(Q(v(t), o(t)))) \quad (3)$$

で決定される.

4.2 Q 学習による注文の組合せ決定

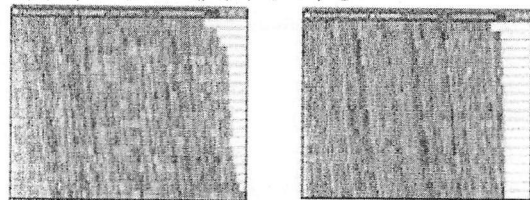
小さな集合で得られた Q 学習則の学習則が, 大きな集合に対しても成立する事が渡辺により確認されている. 本研究では, この経験則を利用して, 以下のように Q 学習を注文の組合せ規則獲得に利用する.

Km 個 ($K < nm$) の注文を nm 個の注文から抜き出す.

- 1) $p=0$ とする.
- 2) ランダムに注文書の順序を生成する. この時, 注文書を m 個づつ分割し, S を生成した後, LCO で $F(S)$ を計算する.
- 3) $p=p+1$ とする. ランダムに r 番目の注文書を選択し, $r-1$ 番目, r 番目, $r+1$ 番目の類似度平均を状態として $X(\theta)$ の計算をし, $x(p)$ から式(4)で状態を $v(\theta)$ とする.

$$v(p) = \{v_1, v_2, \dots, v_c\} \quad v_y = x_y \in [y/c, (y+1)/c] \quad (4)$$

- 4) 式(3)から行動 $o(\theta)$ を決定し, 新しい順序列 S' に LCO で $F(S')$ を計算する.
 - 5) $F(S') < F(S)$ であれば, 式(2)で Q 値を更新する.
- 上記アルゴリズムで, 行動 $o(\theta)$ は, r 番目の注文書を em ($e=1, 2, \dots, K$) 個右に移動するとする.



(a) Initial solution (b) Obtained solution

Fig. 1 Gantt charts applied by LCO (20 storage racks)

5. おわりに

配送問題に LCO と強化学習を適応する方法を提案した. 実際の適用に当たっては, 4.2 章で得られた規則を, 注文列をより大きく区切って S を作成し, 適用すればよい. なお, LCO のみによる結果を図 1 に示した.

参考文献

- 1) 中村 太輔, 渡辺 美知子, 古川 正志: 物流配送計画への LCO の適応, 2004 年度精密工学会北海道支部学術講演会講演論文集, (2004)65
- 2) 松村 有祐, 渡辺 美知子, 古川 正志: ジグ・工具の交換を考慮した JSP の LCO による解法, 2004 年度精密工学会北海道支部学術講演会講演論文集, (2004)63
- 3) 宮川 有樹, 渡辺 美知子, 古川 正志, 木下 正博: 強化学習によるブロック形状自律エージェントのトラック・ヤード抜け出し行動の獲得, 2001 年度精密工学会春期大会学術講演会講演論文集, (2001) 338