

アンツルーティングを活用した配車配送計画手法の効果分析*

Effects of Vehicle Routing and Scheduling Procedures with Ants Routing*

谷口栄一**・山田忠史***・川本宗由****

By Eiichi TANIGUCHI**・Tadashi YAMADA***・Muneyuki KAWAMOTO****

1. はじめに

近年，都市内における交通状況の悪化が著しい．その原因の1つとして，トラックによる貨物輸送が挙げられる．消費者ニーズの多様化に伴う Just-in-Time 輸送をはじめとする高度な物流サービスの要求や，e-コマースの普及が，多頻度・小口輸送を増加させており，その結果，低積載率のトラックが多く用いられている．そのため，都市内における交通渋滞や環境問題の悪化に拍車がかかっており，一方で物流効率も低下するといった悪循環が生じている．

本研究では，これらの問題の解決を図るべく，高度な配車配送計画システムの導入効果に注目する．具体的には，企業（荷主，物流業者など）側の配車配送計画の情報化による集配送費用の削減を目的として，配車配送時の経路選択において，学習アルゴリズムの一つであるアンツルーティングを活用する．そして，このような配車配送システムを導入した場合の効果について，他の配車配送計画手法と比較することにより検討する．

2. 配車配送計画

(1) 定式化

本研究では，都市内の集荷活動に注目し，最適な配車配送計画とは，総物流費用（すなわち，固定費，走行費，遅刻・早着ペナルティの総和）を最小化するように，顧客の割当・訪問順序・出発時刻などを算定するものとする．したがって，配車配送計画問題は，以下のように定式化できる．（制約条件については参考文献1）参照）

*キーワード：物流計画，交通情報，ITS

**フェロー，工博，京都大学大学院工学研究科
（京都府左京区吉田本町，
TEL075-753-4789，FAX075-752-5303）

***正会員，工博，京都大学大学院工学研究科
（京都府左京区吉田本町，
TEL075-753-4787，FAX075-752-5303）

****学生員，工修，京都大学大学院工学研究科
（京都府左京区吉田本町，
TEL075-753-4788，FAX075-752-5303）

Minimize

$$C(t_0, X) = \sum_{k=1}^b \sum_{l=1}^{m(k)} C_{f,l} \cdot \delta(x_l) + \sum_{k=1}^b \sum_{l=1}^{m(k)} C_{t,l}(t_{l,0}, x_l) + \sum_{k=1}^b \sum_{l=1}^{m(k)} C_{p,l}(t_{l,0}, x_l) \quad (1)$$

ただし，

$C(t_0, X)$ ：総物流費用(円)

b ：デポの総数

t_0 ：トラック l のデポ出発時刻を表すベクトル

x_l ：トラック l の配送ルートへの顧客の割り当てと訪問順序を示す数列

$m(k)$ ：デポ k で使用可能なトラック台数の上限

$C_{f,l}$ ：トラック l の固定費用（円/台）

$\delta_l(x_l)$ ：=1; トラック l を使用する時
=0; その他の場合

$C_{t,l}(t_{l,0}, x_l)$ ：トラック l の運行費用（円）

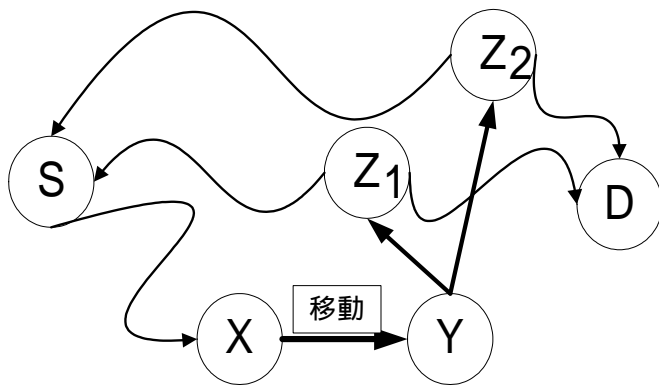
$C_{p,l}(t_{l,0}, x_l)$ ：トラック l の早着・遅刻ペナルティ（円）

(2) モデルの概要

上述の配車配送計画問題は，リンク所要時間を単一値で捉えた配車配送計画モデル(VRPTW-F)に相当する．既存の VRPTW-F モデル¹⁾では，リンク所要時間の平均値に着目した上で，経路探索手法にダイクストラ法が用いられているが，本研究では，リンク所要時間の変動を学習した上で経路探索を行うアンツルーティングを適用する．

本研究では，これら2種類のVRPTW-Fモデルの他に，学習アルゴリズムを内包した経路探索の代表的な手法であるQ-ルーティングを用いたVRPTW-Fモデルも比較対象として取り上げる．さらに，リンク所要時間の変動に伴う早着・遅刻リスクを考慮した確率的配車配送計画モデル(VRPTW-P)²⁾についても，比較対象として採用する．

既存のVRPTW-Fモデルでは，交通シミュレーションによって得られた30日分のリンク所要時間を平均して，単一のリンク所要時間を算出してダイクストラ



目的ノード S	現在ノード Y	
	次の行き先	確率
	X	
	Y	
	:	
	Zi	

このネットワーク内の S で発生したアリが、X から Y に移動したとき確率テーブル更新、つまり学習が行われる。このとき更新されるのはノード Y における確率テーブルであり、S を目的ノードとした上での Y から X へ行く確率である。

図 - 1 アンツルーティングの学習方法

法によって経路探索が行われる。このとき、最短経路は、リンク所要時間を用いて 1 分毎に探索される。一方、アンツルーティングを用いた VRPTW-F モデルでは、30 日分のリンク所要時間データを 1 分毎に学習した上で経路探索が行われる。

(3) アンツルーティング

アンツルーティングは、強化学習を経路選択に応用した手法の一つであり、アリの採餌行動モデルを利用した経路探索手法に相当する。アリが化学物質（フェロモン）によって巣と餌の間の最短経路を発見する方法を応用して、ノード間の最短経路を見つける。この手法は確率的に良好な経路を探索するために、ある特定の状況の下で、ネットワークの変化への適応力に優れている。

アンツルーティングの学習方法を図 - 1 に示す。アンツルーティングでは、ネットワーク内を多数のアリが動き回ることによって、適切な経路が獲得される。

各ノードは、ある宛先を持った上で、次に向かうべきノードを決定するための選択確率テーブルを有しており、アリがノードの移動を行う度に学習される。その学習は、backward exploration と呼ばれる手法で行われる。この方法は、アリが通ってきた経路を遡って学習するものである。つまり、ノード X からノード Y にアリが移動したとき、更新されるのはノード Y にいる時にノード X を次の宛先を選ぶ確率であり、ノード Y の確率テーブルが更新される。確率テーブルの確率値は、以下のような学習式で更新される。

$$\Delta p = k / f(t) \quad (\text{ここに、} k : \text{学習率(分)}) \quad (2)$$

$$P_Y(S, X) = \frac{P_Y(S, X) + \Delta p}{1 + \Delta p} \quad (3)$$

$$P_Y(S, Z) = \frac{P_Y(S, Z)}{1 + \Delta p},$$

$$Z \in \text{neighbor of } Y, Z \neq X \quad (4)$$

式(2)の p は確率値の変化量を表しており、学習率 k をトラベルタイム $f(t)$ で除したものである。トラベルタイムとは、 $X - Y$ 間のリンク所要時間ではなく、アリがノード S で発生したとすると、ノード S で発生してから現在のノード Y に移動するまでの全所要時間である。この変化量の大小によって、学習の進行に伴って確率の大小が変化する。また、学習率 k の大きさによって、学習の速度が規定される。

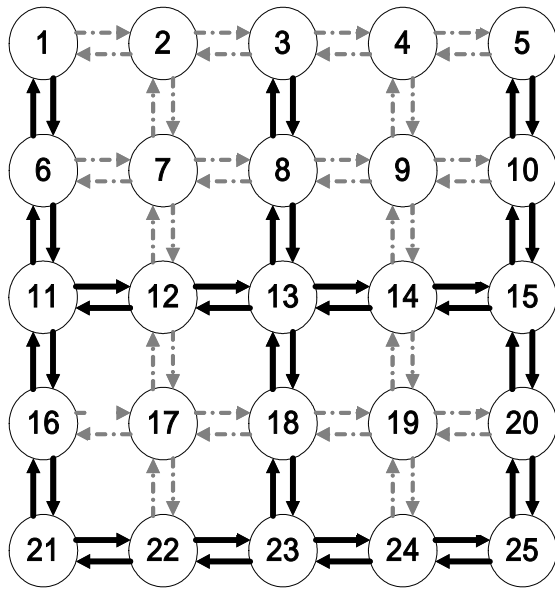
式(3)は、ノード Y での確率更新を表している。ここで更新される $P_Y(S, X)$ は、 S を宛先として、ノード Y から以前にいたノード X へ行く確率を意味する。

式(4)の $P_Y(S, Z)$ は、ノード Y での X 以外の行き先の確率であり、この式によって正規化される。

3. ケーススタディ

(1) 問題設定

図 - 2 に、本研究で用いた仮想道路ネットワークを示す。デポに待機しているトラックの数、種類、ならびにネットワーク上の顧客の位置・需要量は既知とする。物流事業者は 1 社とし、デポの位置はノード 5 に設定する。使用可能なトラックは、デポに 2t, 4t, 10t トラック各 2 台ずつである。顧客の配置については、デポ以外のノードから選び、顧客数を 24 個とし



13 : ノード

← : リンク (自由走行速度 : 20km/h、自由走行時間12分、2車線)

←--- : リンク (自由走行速度 : 15km/h、自由走行時間16分、2車線)

図 - 2 対象道路ネットワーク

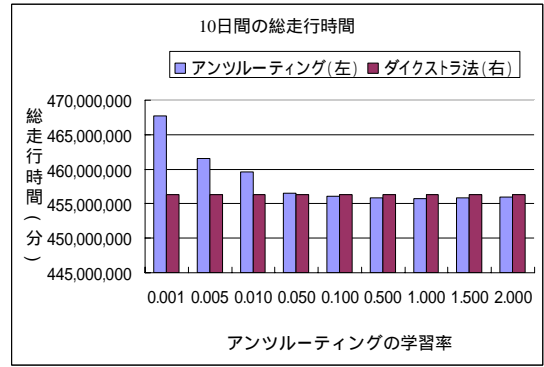


図 - 3 総走行時間の比較 (10日間)

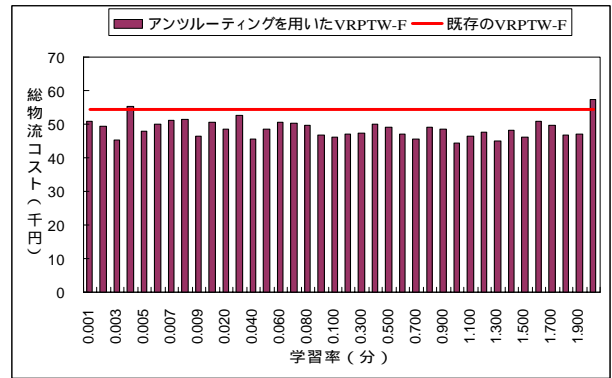


図 - 4 総物流費用の比較

た。

また、顧客は全て集荷先であると仮定し、集荷量は113.5~138 kg、作業時間は10分とした。また、遅刻ペナルティについても全ての顧客で同一とし、438.5 (円/分) に設定した。なお、到着時刻制約は9:00~18:00の時間帯でランダムに設定した。

(2) 既存モデルとの比較

はじめに、30日間の学習を基にして出力された各ODペア間の経路を用いて、その後10日間の走行を行ったときの総走行時間を調べる。図-3は、ダイクストラ法による10日間の各ODペア間の総走行時間と、学習率0.001~2.0の範囲でのアンツルレーティングによる10日間の各ODペア間の総走行時間を比較した結果である。学習率0.001の場合、ダイクストラ法よりも総走行時間が2.5%程度大きくなっている。しかし、学習率が上昇するにつれて、両者の総走行時間の差は次第に小さくなり、学習率が0.1を超えたあたりで、アンツルレーティングを用いた場合の総走行時間がダイクストラ法を用いた場合の総走行時間よりも減少する。以降、この傾向は、学習率2.0まで続く。適切な学習率を用いれば、アンツルレーティングを用いた経路選択が有用であることが確認された。

次に、両者を費用の面で比較する。図-4より、学習率0.001~2.0のアンツルレーティングを使用した

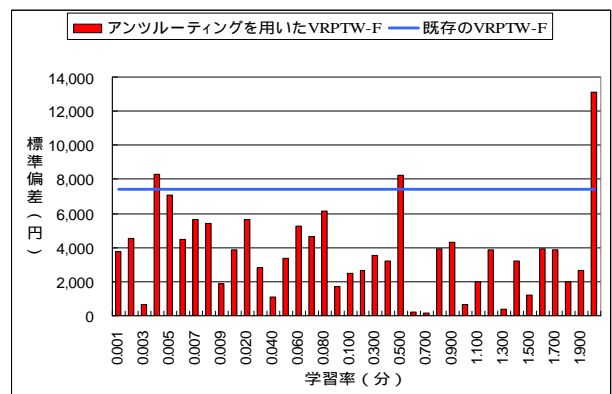


図 - 5 物流費用の標準偏差の比較

VRPTW-Fモデルでは、多くの学習率において、総物流費用は、ダイクストラ法を適用した既存のVRPTW-Fモデルに比べて減少している。費用削減量が最大の場合(学習率1.0)、約18%の費用抑制効果が確認された。

10日間の物流費用の標準偏差を比較した図-5から、多くの学習率において、アンツルレーティングを用いたVRPTW-Fモデルから得られる物流費用の標準偏差は、既存のVRPTW-Fモデルに比べて減少していることが確認できる。学習率が0.001~0.5までは、標準偏差の大小関係が不安定であるが、学習率が0.6~1.9までは安定して、小さな標準偏差を示している。特に学習率0.6と0.7では、標準偏差は約200円と非常に小さな

値を示しており、既存の VRPTW-F モデルから得られる標準偏差（7436 円）を大幅に下回る結果となった。よって、リンク所要時間が変動するような状況下で集配送を行った場合、アンツルーティングを用いた VRPTW-F モデルの適用により、費用面で安定した集配送が実現できるものと考えられる。

（3）他の配車配送計画手法との比較

アンツルーティングを用いた VRPTW-F モデルのパフォーマンスをさらに検証するために、同様の問題を用いて、以下の2つの配車配送計画モデルから得られる結果と比較する。1つは、強化学習手法の中でも基本的な Q 学習を経路探索に応用した Q-ルーティングを使用した VRPTW-F モデルであり、もう1つは確率論的配車配送計画モデル（VRPTW-P）である。

表 - 1 に、それぞれの配車配送計画手法を適用した場合の各費用が比較されている。この結果から、総物流費用が、「VRPTW-P モデル<アンツルーティングを用いた VRPTW-F モデル<既存の VRPTW-F モデル<Q-ルーティングを用いた VRPTW-F モデル」の順で増加している。費用の内訳に注目すると、どの手法においても同種のトラックを2台運行しているため、固定費用は同じになる。したがって、総物流費用の差は運行費用、遅刻ペナルティ、早着ペナルティによるものであり、とりわけ、遅刻ペナルティの差が総物流費用に大きな影響をもたらしていることがわかる。

総費用が「既存の VRPTW-F モデル<アンツルーティングを用いた VRPTW-F モデル」という順になる原因は、主に遅刻ペナルティの減少によるものである。30 日分のリンク所要時間の変動を学習するアンツルーティングが、30 日分のリンク所要時間の平均値を用いて最短経路探索を行うダイクストラ法よりも、その後の 10 日間の集荷において時間の変動を考慮できるものと推察される。また、「既存の VRPTW-F モデル<Q-ルーティングを用いた VRPTW-F モデル」という順になる原因も、主に遅刻ペナルティの減少によるものである。Q-ルーティングでは経路探索が局所解に陥り易く、30 日分のリンク所要時間の変動を十分に考慮しきれなかったものと考えられる。この結果は、適切な学習方法が適用されなければ、学習アルゴリズムを配車配送計画手法に内包しても、リンク所要時間の変動に十分に対処しきれない可能性を示唆している。

総費用が「VRPTW-P モデル<アンツルーティングを用いた VRPTW-F モデル」となる原因は、運行費用と遅刻ペナルティの減少によるものである。遅刻ペナルティの減少は、総費用算定時の遅刻ペナルティの取り扱いの相違に起因するものと考えられる。どちらのモデルも総物流費用最小化を最適化基準としているが、

表 - 1 各費用の比較

単位(円)	VRPTW-P	VRPTW-F	アンツルーティング	Q-ルーティング
固定費用	20835	20835	20835	20835
運行費用	12627	13778	13734	13346
遅刻ペナルティ	0	3771	1710	10743
早着ペナルティ	17	573	0	158
総物流費用	33479	38957	36279	45082

VRPTW-P モデルでは、リンク所要時間の変動に伴う遅刻リスクを考慮するために、VRPTW-F を適用する他の手法に比べて、遅刻ペナルティの発生が抑制される。アンツルーティングが学習するのはリンク所要時間のみであり、訪問先での到着時刻制約を考慮した上での経路探索にはなっていない。したがって、10 日間の集荷において、アンツルーティングを用いた VRPTW-F モデルでは、遅刻ペナルティが発生した。また、運行費用についても VRPTW-P モデルが下回った結果には、本研究で用いた問題設定が関係しているものと考えられる。10 日間の集荷において、全体的にリンク所要時間が増加するような設定であったために、時間変動の少ないリンクを選択する傾向にある VRPTW-P モデルの方が、運行費用を抑制できたものと推測される。ただし、この点については、多様な問題設定下での結果比較や、より有力な学習率の設定などによって、今後も検討する必要がある。なお、計算時間については、アンツルーティングを用いた VRPTW-F モデルの方が、VRPTW-P モデルよりも約 5 分の 1 であったことから、アンツルーティングを用いた VRPTW-F モデルの有用性の一端が確認された。

5. 結論

本研究では、強化学習手法を経路選択に応用したアンツルーティングに注目し、それを既存の配車配送計画手法(VRPTW-F)に取り入れることで、リンク所要時間の変動を考慮した配車配送計画手法の効果を検討した。都市内の集荷を想定した比較分析の結果、適切な学習方法や学習率が採用されれば、既存の配車配送計画手法と比較して、総物流費用の抑制に効果的であることが示された。

参考文献

- 1) 谷口栄一, 山田忠史, 細川貴志: 都市内集配送トラックの配車配送計画の高度化・共同化による道路交通への影響分析, 土木学会論文集, No.625/ -44, pp.149-159, 1999.
- 2) 谷口栄一, 山田忠史, 柿本恭志: 所要時間の不確実性を考慮した都市内集配トラックの確率論的配車配送計画, 土木学会論文集, No.674/ -51, pp.49-61, 2001.
- 3) D. Subramanian, P. Druschel and J. Chen: Ants and reinforcement learning: A case study in routing in dynamic network, Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-97), pp832-839, 1997.