

# マルチエージェント深層強化学習を用いた大規模都市型水害後における企業の行動制御

楊少鋒・小川芳樹・池内幸司・秋山祐樹・柴崎亮介

## Firm-level behavior control after large-scale urban flooding using multi-agent deep reinforcement learning

Shaofeng YANG, Yoshiki OGAWA, Koji IKEUCHI,  
Yuki AKIYAMA and Ryosuke SHIBASAKI

**Abstract:** With natural disasters have become large scale, diversified, and frequent, the indirect economic damage due to interruption of supply chain tend to be large. Therefore, it is important to recovery as quick as possible for companies after disasters. In this paper, for the purpose of the strategic action of companies for efficient recovery of supply chain, we focus on large-scale flood in Tokyo metropolitan, and build a simulation environment for the process of spillover of economic damage and recovery in supply chain with combination of multiple environment big data such as flood simulation analysis data, large-scale inter-firm transaction data and road data. Next, we propose an action policy for companies based on multi-agent deep reinforcement learning to optimize the behavior of companies in the recovery process.

**Keywords:** 強化学習(reinforcement Learning) , 企業間取引データ(inter-firm transaction data) , サプライチェーンマネジメント(supply chain management) , 氾濫解析 (flood simulation) , 都市型水害(Urban flooding)

### 1. はじめに

近年、記録的な豪雨や地震による企業の操業停止によってサプライチェーンの寸断が多発し、企業間取引を通じて経済に甚大な影響を及ぼした。これに対して、災害時の早期復旧を目指す事業継続計画(BCP)を立案することで被害を最小限に抑えることが可能である。しかし、東京商工リサーチのアンケート調査(2018)によると、日本企業の

BCP策定率は27.7%であり、特に水害を意識したBCPを行っているのはわずか全体の0.22%であり、欧米諸国と比較して非常に低い水準である。これはBCPを策定するためのデータの整備がなされていないことが原因の一つであると考えられる。そのためサプライチェーンのインパクト推計や復興過程をシミュレーションすることは困難である。

サプライチェーンにおける負のインパクトの波及について、Goto et al. (2017) は、大規模企業間取引データを用いて、モンテカルロの手法で企業間の倒産連鎖のメカニズムについて分析を行っ

---

楊少鋒 〒227-8563 千葉県柏市柏の葉 5-1-5

東京大学新領域創成科学研究科

Phone: 04-7136-4003

E-mail: sfyang@csis.u-tokyo.ac.jp

ている。Ogawa et al. (2019) は大規模企業間取引データや津波推定結果を用いて南海トラフ巨大地震による経済被害がどのようにサプライチェーンに波及し、復興していくかについて分析している。しかし、サプライチェーンの個別企業単位での被害波及から復興までのメカニズムまで踏み込んだ定量的な評価手法については、明確な解答がない。

一方、近年では実世界の大规模システムの分析にマルチエージェントシミュレーションが多く使用されている。また、機械学習手法を用いて大规模データから統計的関係性を学習した結果をもとに、様々な分野で予測や分類を行う研究が多く行われている。大规模で複雑なシステム構造を持つサプライチェーンの分析には、機械学習手法が適していると考えられる。

以上を踏まえて、本研究では、荒川が決壊することを想定した氾濫解析結果データや大规模企業間取引データを使用し、災害後発生から復興までの意思決定過程について、個々の企業を考慮したサプライチェーンの挙動をマルチエージェントモデルを用いて構築する。また、機械学習手法を用いて復興過程において効率的な復興を目的とした個別企業の行動戦略を得るための学習フレームワークを提案する。

## 2. 使用データ

### 2.1 荒川氾濫解析データ

本研究では、荒川の氾濫による水害を想定する。

国土交通省荒川河川事務所が保有する、荒川流域の浸水想定区域図作成の際に行われた国土交通省荒川河川事務所の氾濫シミュレーション結果を用いた。対象洪水は年超過確率 1/1000 の規模の洪水である。データセットには堤防の決壊所が異なる 335 のシナリオのデータが格納されており、各シナリオの決壊場所、100 m メッシュごとの浸水深の時系列データ(決壊してから 1 か月間の 10 分単位のデータ)などの情報が格納されている。本研究で使用するデータは右岸 21km 地点に想定堤防決壊箇所があるシナリオを使用する。(図 1)

### 2.2 企業本社データ、及び企業間取引データ

株式会社帝国データバンクが保有する 2016 年の約 500 万件の企業の本社間の取引データを使用した。それぞれの取引は、一年間で一回分のデータである。企業 ID によって、企業及び事業所データと関連づけるとでネットワーク分析を行うことが可能になる。データには取引の受注社と発注社の企業コード、取引品目、推定取引金額などの企業の本社間の取引情報が格納されている

## 3. 提案手法

### 3.1 マルチエージェント深層強化学習

深層強化学習は強化学習と深層学習と組み合わせた手法である。強化学習では、環境との相互作用を通して、自身の行動ポリシーの最適化を自ら行っていく。時間  $t$  において、エージェントは状態  $S_t$  を観測する。次に制御指針の集合  $A$  から行動  $a_t (a_t \in A)$  を実行して得られる報酬  $R_{S_t, S_{t+1}}^{a_t}$  を用い



図-1 浸水の拡大状況 (想定堤防決壊箇所：右岸 21.0km)

て状態 $S_t$ における行動 $a_t$ の行動価値関数 $Q(S_t, a_t)$ を式1の通り更新を行う。

$$Q(S_t, a_t) \leftarrow Q(S_t, a_t) + a \left[ R_{S_t, S_{t+1}}^{a_t} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, a_t) \right] \quad (式1)$$

$a(0 \leq a \leq 1)$ は学習率を示し、 $\gamma(0 \leq \gamma \leq 1)$ は報酬の割引率である。深層強化学習は、 $Q(S_t, a_t)$ が膨大になる場合、多層ニューラルネットワークを用いて $Q(S_t, a_t)$ を近似する方法である。

### 3.2 状態空間と行動の設計

被災地である荒川流域に立地する約3万の企業とその取引先約17万の企業エージェントとする。エージェント同士は互いに取引先としてネットワークで繋がることできる。被災時より1stepを一日とし、2ヶ月かけて復興するとして、60stepシミュレーションを行う。

被災エージェントは取引先を失うとし、復興するために以下の行動を取ることができるとする。

「復興作業」、「過去の取引先と契約する」、「新取引先を開拓する」、「事業拡大」、「行動せず」の5種類とする。なお、取引先の選定方法に取引の向きや距離を分け、合計9パターンの行動を取ることができるとする。

また、企業エージェントの行動は他エージェントへ影響を与えることができるとする。取引先になると互いに受注と発注の関係を持ち、生産活動に影響を与える。また、エージェント間での利害関係を考慮した行動を学習できるようにするため、Foerster et al. (2016)が考案したエージェント間の相互作用を考慮したマルチエージェント強化学習である Reinforcement inter-agent learning (RIAL)を使用する。具体的には、図2の通り、エージェント1は、時間 $t$ において、環境の観測値 $o_t$ とエージェント2との相互作用 $m_{t-1}$ の両方を学習データとして学習関数であるQ-Netに取り込み計算を行う。その後、エージェント2も同様に学習する。これによりエージェント間の相互作用を考慮した行動を学習することができ、自社の利益

最大化だけでなく、エージェント同士の利益を考慮した行動を取ることができるようになる。

シミュレーションのゴールをサプライチェーン全体の取引金額が被災前の8割の水準に戻るまで

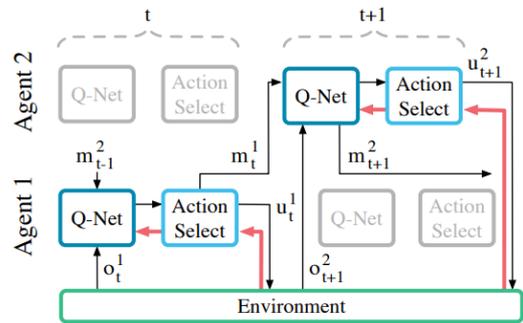


図-2 RIAL のアルゴリズム

とする。なお、復興期間中の企業の生産力はコブ・ダグラス型生産関数(式2)に従うとする。本研究では、スケール係数 $A$ を固定とし、生産額 $Y$ は生産要素である労働力 $L$ の労働分配率 $\alpha$ 乗に資本 $K$ の資本分配率 $\beta$ 乗を乗じて求めるとする。

$$Y = AL^\alpha K^\beta \quad (式2)$$

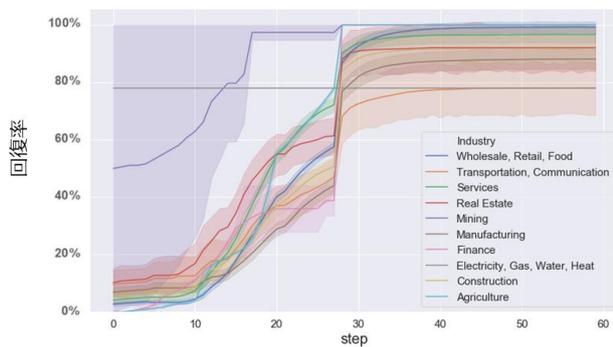
$s.t. \alpha = 1 - \beta, 1 > \alpha > 0, 1 > \beta > 0.$

本研究では、企業の生産額の計算条件を以下のように設定する。労働力 $l$ を従業員とし、資本 $k$ をライフラインと建物の被害で評価する。ライフラインは道路の浸水状況 $r$ と電気設備 $e$ である道路と電気設備 $e$ とする。通常時の生産額を $y$ とした時、水害による被災企業の生産減少額 $\Delta y$ は労働減少率 $\Delta l$ (=従業員の死亡率)と資本減少率 $\Delta e \cdot \Delta r$ (=電気設備被害・道路の浸水状況)から式3のように求めるとする

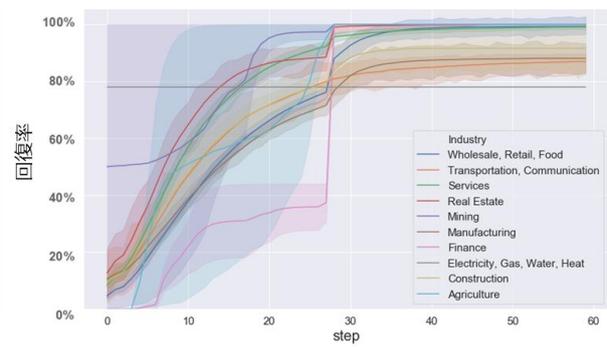
$$\Delta y = \left\{ \alpha \cdot \frac{\Delta l}{l} + (1 - \alpha) \left( \frac{\Delta e}{e} \cdot \frac{\Delta r}{r} \right) \right\} * y. \quad (式3)$$

### 4. 学習結果及び評価実験

図3は学習前(図-3a)と学習済(図-3b)のエージェントの行動によるサプライチェーンの復興状況の産業別での比較である。学習前(ランダム行動)の場合、被災直後は効率的に復興できず、一定期間は無駄な行動をとるエージェントが多く存在する。10 Step目になると徐々に復興し始める。一方、学習済では被災直後から多くの企業が有意義な行動を取っており、未学習モデルと比較して早い段階で回復していることが確認できる。



(a) ランダム行動, 復興状況



(b) 学習済, 復興状況

図-3 学習前後の比較(a):ランダムに行動した場合の復興状況 (b):ランダムに行動した場合行動の推移

図4は学習済みモデルを用いて200回シミュレーションした結果の平均回復率のヒストグラムである。平均で75% - 80%回復することができることがわかった。また、結果の分布は正規分布に従うと考えられる。

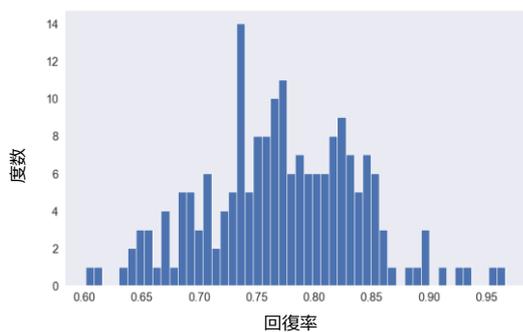


図-4 学習済みモデルの信頼度検証 (n=200)

## 5. 終わりに

本研究では、マルチエージェント深層強化学習の手法を用いて、災害後サプライチェーンの企業間取引の復興過程において、企業の行動戦略の最適化を行う手法を提案した。災害の被害推定結果及び大規模事業所間ネットワークデータを組み合わせることで、災害の被害波及から復旧過程のシミュレーション環境を構築した。また個々の企業をエージェントとした発災から復興までの行動戦略を得るための強化学習フレームワークを設計し、東京都が大規模洪水によって被災した場合をケーススタディとして、実証実験を行った。発災段階から復興に至るまでの各々の復興段階に応じた適切な企業行動（設備投資・代替取引など）も確認することができた。

今後の課題として、2つの点が挙げられる。1つ

目はフレームワークの強化である。より正確な経済的影響の推定をするために、シミュレーションの各パラメータ設定、企業間の相互作用、強化学習手法など十分に考慮できていない部分に関しては今後改良していく必要がある。2つ目は、現在BCPを策定している企業の復旧計画や災害対策の状況をモデルに取り込み、新たなBCP対策について、機械学習を用いた評価手法の確立を目指したい。

## 謝辞

本研究は国土交通省荒川河川事務所及び、株式会社帝国データバンクよりデータ提供を受けて実施したものである。また、三菱総合研究所の大熊裕輝様より建設的なアドバイスを頂いた。ここに記して謝意を表したい。

## 参考文献

- 東京商工会議所, 2018, 会員企業の防災対策に関するアンケート. 2019/8/31 最終アクセス.  
 <<http://www.tokyo-cci.or.jp/page.jsp?id=1000399>>  
 Foerster, J. N. Assael, Y. M. de Freitas, N. and Whiteson, 2016. S. Learning to communicate with deep multi-agent reinforcement learning. CoRR, abs/1605.06676.  
 Goto H, Takayasu H, Takayasu M, 2017. Estimating risk propagation between interacting firms on inter-firm complex network. PLoS ONE 12 (10): e0185712. <https://doi.org/10.1371/journal.pone.0185712>  
 Ogawa, Y., Akiyama, Y., Yokomatsu, M., Sekimoto, Y., Shibasaki, R., 2019. Estimation of supply chain network disruption of companies across the country affected by the Nankai Trough Earthquake Tsunami in Kochi City, Journal of Disaster Research, Vol. 14, No. 3.