

深層強化学習を用いたサプライチェーン復興過程の最適化手法の提案 - 2016年熊本地震を事例として -

楊少鋒, 小川芳樹, 池内幸司, 柴崎亮介, 大熊裕輝

Optimization Method for Supply Chain Reconstruction Process Using Deep Reinforcement Learning - A Case Study of the 2016 Kumamoto Earthquake -

Shaofeng Yang*, Yoshiki Ogawa, Koji Ikeuchi, Ryosuke Shibasaki, Yuuki Okuma

Abstract: To improve the resilience of supply chains (SC), it is necessary for individual companies to take optimal actions autonomously. In this study, using the Kumamoto earthquake as a case study, we construct a multi-agent simulation environment of the recovery process of a firm-based SC using large-scale inter-firm transaction data. Next, we construct a model of corporate behavior based on game theory that takes into account the interactions among the firms, and use deep reinforcement learning to train the corporate agents to autonomously take optimal actions.

Keywords: マルチエージェントシミュレーション(multi-agent simulation), 深層強化学習(deep reinforcement learning), サプライチェーンマネジメント(supply chain management), 企業間取引データ(inter-firm transaction data), ゲーム理論(game theory)

1. はじめに

近年, 世界における自然災害は大規模化, 多様化し, 頻発するようになり, 今後も増加する傾向があると予測されている. 自然災害は人的, 物的被害だけでなく, 経済に対しても大きなインパクトを与える. そのため被災地の企業が被災することによって企業間の取引ネットワークを通じて間接的に多くの企業が生産停止するなどの被災地以外の企業も影響を受ける.

例えば 2011 年に起きたタイ北部チャオプラヤ川の水害において, 川沿いの工業団地が浸水し, 日本企業を含む 800 社を超える企業や工場が被災した. 当時, タイは世界第二位のハードディスクの生産国であり, 世界の生産量の約 5 割が集中している. この災害により, サプライチェーンを通じて世界的にハードディスクの品薄状態が起こり, パソコン, 録画再生機など電子機器の製造業界に大きく影響を与えた. 地震災害においても同様にボトルネック企業の被災によるサプライチェーンの寸断がみられた(三菱UFJリサーチ&コンサルティング株式会社 2020). このように世界的によるサプライチェーン寸断による経済的な影響が多いいため, 世界的にサプライチェーンのレジリエンスの向上が喫緊の課題となっている. 日本では, 経済産業省は令和 3 年度経済産業政策の重点取組の一つに「サプライチェーン強化・サプライネットの構築」を掲げている. ま

た, 2020 年にサプライチェーンの途絶によるリスクが大きい製品の保全を目的とした補助金に 1 兆 7 千億円の応募が集まった. しかし, 多くの企業は災害後にどのような行動をとるべきかについて, 検討されていない, または検討するために必要な材料が不足しているという課題がある. 災害や事故が発生したときにその企業の生産活動が中断しないような, あるいは万一中断しても目標復旧時間内に再開させる事業継続計画 (BCP) を策定するようにガイドラインが策定されている. 東京商工会議所のアンケート (2018) によると, サプライチェーンの確保を行っている企業は, 会員企業全体の 1.1%(1127 社中の 12 社)であり, 非常に低水準である.

以上を踏まえて, 各企業が最適な行動を取り, 災害などサプライチェーンに対して迅速に回復できるようなインテリジェントなサプライチェーンを構築するための研究手法が必要である.

一方で海外においては, アメリカのコンサルティングファーム Mckinsey & Company が「Resetting supply chains for the next normal」の中でサプライチェーンの再構築の重要性を指摘している. 中国においては, 2021 年の中央経済工作会議で「産業チェーン・サプライチェーンの自主コントロール能力を強化する」を挙げ, サプライチェーンの自身が強い危機対応能力を持って様々な負のインパクトに対応していく必

* 学生会員 東京大学新領域創成科学研究科
〒227-8563 千葉県柏市柏の葉 5-1-5 Tel : 04-7136-4003 または E-mail : sfyang@csis.u-tokyo.ac.jp

要があると指摘している。

SC レジリエンスの向上を目的としたサプライチェーンネットワークデザイン (SCND) に関する既往研究は数多く報告されている。SCND の方法論として大きく 3 種類に別れ、1 つ目に最適化 (Optimization) である。2 つ目にシミュレーション (Simulation)。そして 3 つ目はこの 2 つを組み合わせた Optimization-simulation (opt-sim) である。コンピュータ産業の発展によって、近年では sim-opt の手法が盛んに行われている。主に数学的アプローチやオペレーションズ・リサーチの最適化手法組み合わせ、シミュレーションを通じて分析を行っている。また企業活動に関するデータ蓄積を背景にサプライチェーンの影響を広域で分析するために企業間取引ビッグデータを用いた経済被害分析も行われている。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、企業間の倒産連鎖のメカニズムを明らかにするために、企業間取引ビッグデータを用いて、モンテカルロ手法で分析を行った。また、将来的に予測されている大規模自然災害による経済的影響に関する研究の例として、Ogawa et al. (2019) は企業間取引データや津波データなど多様なデータを組み合わせることで日本における南海トラフ地震後における企業の復興過程を推定した。

しかし、既往研究では以下の課題がある。

実際のサプライチェーンデータを用いて復興過程での企業間の相互作用をゲーム論に基づくシミュレーションモデル行っている研究はまだない。ゲーム理論を導入することで、サプライチェーン全体の利益を最大化するための意思決定を最適化することが可能になる。

以上を踏まえて、本研究はサプライチェーンレジリエンスの向上のために、大規模なサプライチェーンデータを用いて、sim-opt 手法によるゲーム理論に基づくサプライチェーンの意思決定支援モデルを提案する。まず、企業間取引データと災害データを用いて、マルチエージェントベースのサプライチェーンシミュレーション環境を構築する。次に、企業が自主的に最適な行動が取得できるように、機械学習及びゲーム理論に基づく意思決定モデルを取り入れる。

最後に、学習済みモデルを用いたシミュレーションを行う。

2. データ

2.1. 企業概要情報データ

民間信用調査企業である株式会社帝国データバンクが保有する 2015 年、2016 年時点の日本企業の企業概要データ約 165 万件を使用する。このデータセットは日本の法人企業 187 万社の約 9 割の規模を有し、日本全国の企業を網羅している。通常は企業の概要情報を取得しづらい中小企業についても網羅性が高い点が特徴である一方、信用取引すなわち企業間取引 (BtoB) を主体とする企業を中心とする点に留意が必要である。データには、企業をユニークに識別することが可能である帝国データバンクが独自に付与した企業コードをはじめ、本社所在地、業種情報、売上高などの各企業の概要情報が格納されている。

2.2 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 2015 年、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。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、取引高推定値などの情報が格納されている。なお、取引高推定値に関しては、株式会社帝国データバンクと東京工業大学高安研究室の共同研究により開発されたものである (Tamura et al., 2018)。図 1 のように、企業 ID によって、企業及び事業所データと関連づけるとしてネットワーク分析を行うことが可能になる。また、企業間の取引データには、モノとお金の流れがあり、取引ネットワークは有向グラフとして評価する。

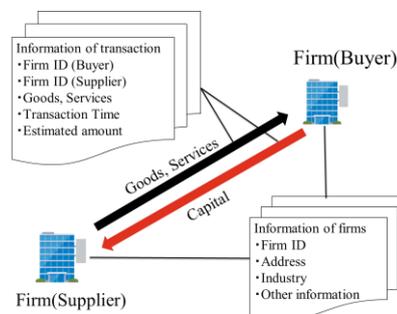


図 1：企業間取引データの概要図

2.3. 2016 年熊本地震測定震度データ

本研究では 2016 熊本地震において、地震後のサプライチェーンの復興過程のケーススタディを行う。図 3 のように、気象庁が観測したデータによると、当該地震において最大計測震度 7 である熊本県を中心として、九州地方全域に渡って広い範囲で強い地震が発生した。当該地域での生産停止がサプライチェーンの途絶をもたらし、生産停止が他地域に波及する例があった。生産停止期間は数日から数週間の企業が多いが、一部の大きな被害を受けた企業は事業再開が長期化する場合もあった。

3. 提案手法

2.1. 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、 $Q(S_t, a_t)$ がに

2.2. マルチエージェントシミュレーション

マルチエージェントシミュレーションは、複数のエージェントに同時進行的に各々のルールをもと、お互いに相互作用を受けながら行動するシミュレーション手法である。本研究では複数の企業を主体とした災害後のサプライチェーンの復興過程のシミュレーション環境を構築する。被災企業とその取引先をエージェントとして、ランダムに 13 の企業を選んだ。エージェント同士は互いに取引先としてネット

ワークを構成している。

被災時より、被災企業の被災度合いによって操業率が減少し、取引金額が減少する。そして、時間の経過とともに、操業率が回復していく。復興期間中において、数ある k_t を時間 t における操業率とし、 d_i を企業 i の被害率とし、 s_{ij} を企業 i と企業 $j \in \{1, \dots, n\}$ の取引金額とすると、企業 i の取引総額 S_i は式 2 で表す：

$$S_i = \sum_j^n s_{ij} \cdot (d_i - (d_i \cdot k_t)) \quad (式 2)$$

なお企業の被害率は東京都総務局総合防災部防災管理課の建物全壊率曲線を用いる。操業率は道下ほか(2016)の企業の操業率曲線を用いて計算する。

2.3. ゲーム理論に基づく学習

本研究では、シミュレーションにおいて全ての企業エージェントの被害が最小になるように、サプライチェーン全体の取引規模の回復度を考慮し、全ての企業が最もこれに近づくような行動の組み合わせを取得することを目的とする。そのため、ナッシュ均衡に基づいて、企業エージェントの行動報酬を設定する。 a_i をエージェント i の行動とすると、 a_{i-1} はエージェント i 以外のエージェントの行動である。式 3 を満たす場合、 a_i^* はナッシュ均衡時の行動である。

$$Q(a_i^*, a_{i-1}^*) \geq \arg \max Q(a_i, a_{i-1}^*) \quad (式 3)$$

行動価値関数に入力する状態関数 S_t = (ステップ数, 全体の平均取引額回復率, 各企業の取引回復率) とする。エージェントの行動 a_t は取引額を増加・減少する金額とする。市場の需要を全ての企業が均等に配分することを考慮し、取引額余分に増加した場合、報酬を減らす。また、自社の取引額を減らし他の企業に配分取引額を減少時間 t の市場の需要を m_t とすると個別の企業エージェントの報酬関数は $r_i(t)$ は式 4 で表す：

$$r_i(t) = -\frac{m_t \cdot a_t}{n} \quad (式 4)$$

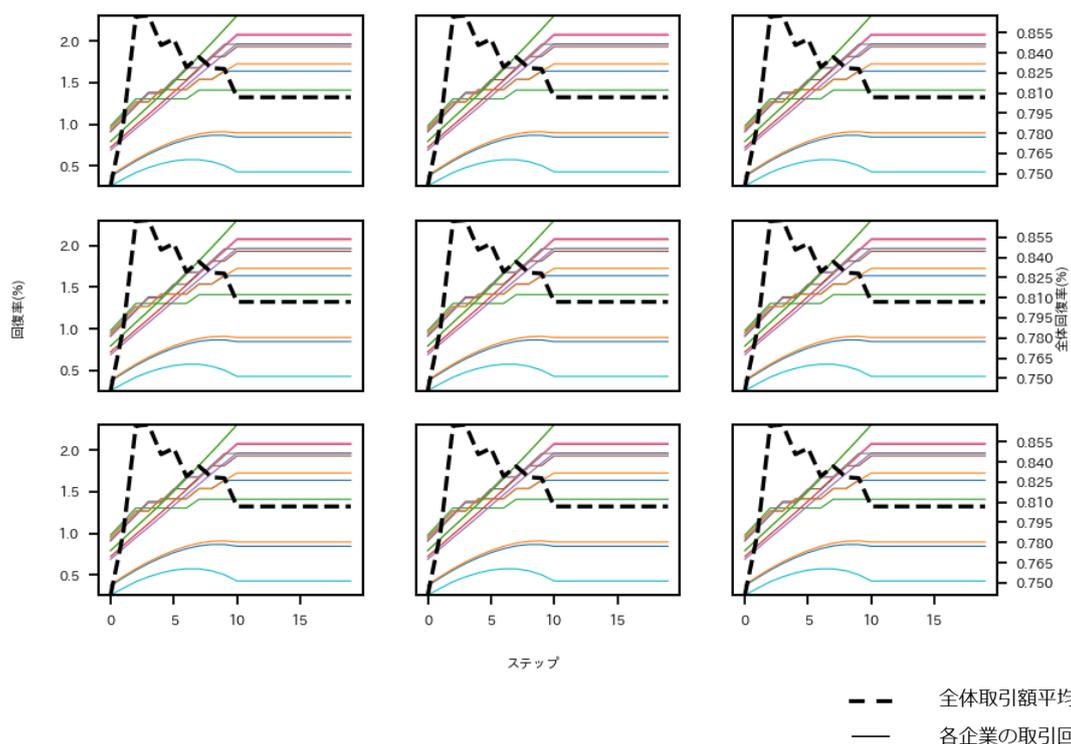


図 2：各企業エージェント回復率及び全企業エージェントの平均回復率の推移

4. 実験結果

図 2 は学習済みモデルを用いて 9 回シミュレーションした結果である。ステップは 20 とし、エージェント数は 13 である。災害前の取引水準を 100% とし、シミュレーション開始後、災害によって取引金額は企業の被災状況によって減少する。その後、各企業の操業率は徐々に回復していき、取引金額を増加する行動を取る。ステップ 5 頃からは一部の企業が取引金額を減少させた。最終的に全体の取引金額の回復率は約 81% に安定した。また、被害が大きい企業は最終的に 100% 回復せずに取引額が減少した。被害が少ない企業が減少分を獲得し、自社の拡大につながった。

5. まとめ

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、ゲーム理論に基づく企業の行動戦略の最適化を行う手法を提案した。

今後の課題として、2つの点が挙げられる。1つ目はシミュレーション環境の強化である。現実では企業の企業の挙動は非常に複雑であり、それを再現するために、経済のモデルを取り入れる必要がある。

2 つ目はモデルの検証である。学習済みのモデルについて、行動価値関数のブラックボックスを解析することで、企業エージェントの挙動をより詳しく把握することが可能である。

謝辞

本研究は株式会社帝国データバンクとの共同研究の研究成果の一部である。

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. 2021/8/31 最終アクセス.
- 道下友哉, 山岸邦彰, 2017, 2016 年熊本地震における事業継続アンケート調査. 日本建築学会.
- 三菱UFJリサーチ&コンサルティング株式会社, 令和 2 年度 産 業 経 済 研 究 委 託 事 業, https://www.meti.go.jp/meti_lib/report/2020FY/000065.pdf 2021/8/31 最終アクセス.
- 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.
- 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>
- Tamura, K., Takayasu, H., Takayasu, M. (2018). Diffusion-localization transition caused by nonlinear transport on complex networks, Scientific Reports 8, Article number:5517.