

令和 6 年 6 月 20 日現在

機関番号：13102

研究種目：基盤研究(B)（一般）

研究期間：2020～2023

課題番号：20H04172

研究課題名（和文）通信データ独占を克服するための疑似データ生成プラットフォームの構築

研究課題名（英文）Development of a Fake Data Generation Platform to Overcome the Monopolization Problem of Communication Data

研究代表者

渡部 康平（WATABE, Kohei）

長岡技術科学大学・工学研究科・准教授

研究者番号：10734733

交付決定額（研究期間全体）：（直接経費） 11,300,000円

研究成果の概要（和文）：本研究課題による成果により、Generative Adversarial Network (GAN)や Variational AutoEncoder (VAE)などに代表される生成系技術を活用し、量的・質的な意味で限られた実データから、実データの代わりに利用可能な疑似データを無限に生成可能な生成器を開発することができた。開発はグラフデータの生成技術とトラフィックデータの生成技術の両面から進められ、特にグラフデータの生成技術においては、既存の生成技術に比べて高い生成性能を実現することができた。上記の成果を2件の論文誌、7件の国際会議、14件の国内会議で成果を発表している。

研究成果の学術的意義や社会的意義

本研究課題による成果は、量的・質的に限られたデータから擬似的な通信データの生成を実現するものであるが、これは近年注目される巨大IT企業によるデータの独占の問題に対して、解決の糸口を提供するものである。また、本研究に関連する技術は、2件の招待講演により発表されており、学会においても注目を集めている。加えて、一流誌であるIEEE Transactions on Network Science and Engineeringやトップカンファレンスである AAAI ICWSM 2023にアクセプトされるなど、学会において高い評価を得ている。

研究成果の概要（英文）：As a result of this research project, we have developed a generator capable of infinitely producing fake data that can be used in place of actual data, by leveraging generative technologies such as Generative Adversarial Networks (GAN) and Variational AutoEncoders (VAE). This development was advanced from both the graph data generation and traffic data generation perspectives. In particular, we achieved superior generation performance in the graph data generation technology compared to existing generation technologies. The outcomes of this work have been presented in 2 journal papers, 7 international conferences, and 14 domestic conferences.

研究分野：通信ネットワーク

キーワード：ネットワーク 生成モデル ニューラルネットワーク 通信データ

1. 研究開始当初の背景

ネットワークに関する製品テストやシミュレーションによる実験を実環境に近い環境で実施するためには、ネットワークの実データ(トラヒックの時系列, フロー構成, トポロジなど)が必要不可欠である. AT&T など米国を中心とする一部の企業や組織は, 保有する実データに基づく実験を行うことで, 様々な論文やイノベーションにつなげている. しかし, それらの企業・組織が保有するネットワークの実データは企業や組織内で利用されるのみで, そのほとんどが公開されておらず, 大学を始めとする大多数の研究者や開発者は, それらのデータにアクセスすることができない. この問題の構造は, 近年騒がれる巨大 IT 企業によるデータ独占と同様であり, 独占により先進的な技術の創出が制約を受けている. 多くの研究者や開発者は, 公開されている僅かなデータや自社で保有している限られたデータを元に, 研究開発を行うしかない.

本研究課題では, このデータ独占によるイノベーションの阻害の問題を, 近年, 機械学習分野で注目を集めるデータ生成技術を活用して解決する. データ生成技術を活用し, 実データの代替として利用可能で, かつ実験の利用目的に沿った任意の特性を持つ疑似データを生成する方法を探求することが, 本課題における学術的問いである.

2. 研究の目的

本研究課題の目的は, 僅かな実データを元に, 限りなく実データに特性が近く, かつ任意のパラメータを調整可能な疑似データを無限に生成可能な生成器から成る実験プラットフォームを構築することである. 本研究課題で開発する生成器は, モデル選定とチューニングの困難性を回避しつつ, 量的・質的な意味で限られた実データから無限に疑似データを生成することを可能にする. 生成された疑似データは実データの特性を保存し, モデルの不整合により結果が乖離する問題を解決する. 加えて, 開発する生成器は, 実データの特性を残しつつ, トラヒックであれば, 平均送信レート, 送出期間長分布の減衰, 特定ポートの出現頻度など, トポロジであれば, ノード数, スケールフリー性やクラスターリング係数など任意のパラメータを変化させることが可能である. そのため, 実データとは異なるものの, 実データの多くの特性をそのまま残すことから, 結果の信頼性を保ちつつ, 柔軟な実験を可能にする.

3. 研究の方法

本研究課題では, 以下に示す 6 つの研究項目を, 研究代表者と 2 名の研究分担者で分担して進める.

- トラヒック生成技術の開発
- トポロジ生成技術の開発
- 任意パラメータの調整技術の開発
- 品質モデルとメッセージ履歴への拡張
- トラヒック行列と移動軌跡への拡張
- 統合的なプラットフォームの構築

最初に, 渡部と津川がそれぞれ主導して, 実データと類似した特性を示すトラヒック生成とトポロジ生成を実現する生成器を開発する. その後, 両技術に共通となる任意のパラメータを調整する技術を開発し, 各生成モデルと組み合わせる. さらに, トラヒックとトポロジに関する生成技術をそれぞれ拡張し, 品質モデルやモビリティモデルの構築に展開する. 最後に各生成モデルを統合し, シミュレーション・実験環境を構築するためのプラットフォームを開発する.

4. 研究成果

本研究課題による成果により、Generative Adversarial Network (GAN)や Variational AutoEncoder (VAE)などに代表される生成系技術を活用し、量的・質的な意味で限られた実データから、実データの代わりに利用可能な擬似データを無限に生成可能な生成器を開発することができた。上記の成果を2件の論文誌、7件の国際会議、14件の国内会議で成果を発表している。加えて、本研究に関連する技術は、2件の招待講演により発表されており、学会においても注目を集めている他、研究期間後に更に1件の招待講演と1件の招待サーベイ論文での発表が予定されている。

開発はグラフデータの生成技術とトラヒックデータの生成技術の両面から進められ、下記に挙げる特徴量に関して学習データセットの特性を保持したまま生成データを生成することを確認した。

- グラフデータの生成技術
 - 平均最短経路長
 - クラスタリング係数
 - 次数分布の冪指数
 - 平均次数
 - モジュラリティ
- トラヒックデータの生成技術
 - 平均トラヒックレート
 - トラヒックレートの標準偏差
 - トラヒックの上下降トレンド

さらに、学習データの特性を保持したまま、一部の特性を調整することの可能とする技術も開発した。特にグラフデータの生成技術においては、既存の生成技術に比べて高い精度で生成可能であることを確認しており(図1参照)、IEEE Transactions on Network Science and Engineeringに論文がアクセプトされるなど、国際的に高い評価を得ている。

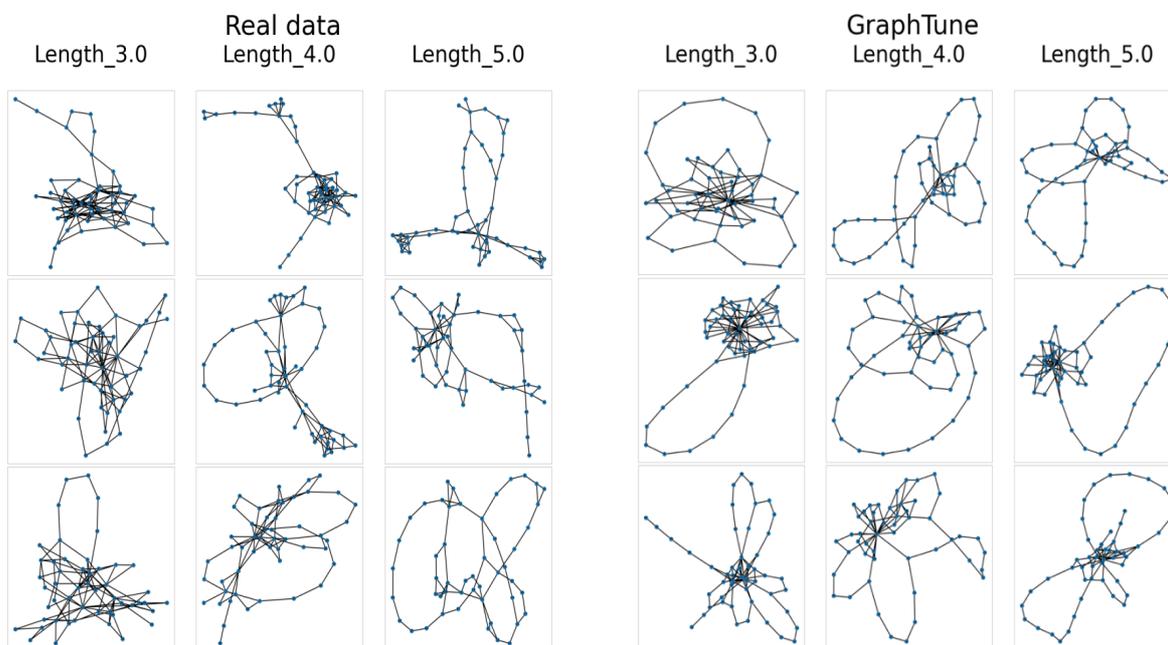


図1 学習データと開発した生成器の出力結果の比較

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 Kohei Watabe, Shohei Nakazawa, Yoshiki Sato, Sho Tsugawa, Kenji Nakagawa	4. 巻 -
2. 論文標題 GraphTune: A Learning-Based Graph Generative Model With Tunable Structural Features	5. 発行年 2023年
3. 雑誌名 IEEE Transactions on Network Science and Engineering	6. 最初と最後の頁 1-13
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TNSE.2023.3244590	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Kohei Watabe, Norinosuke Murai, Shintaro Hirakawa, Kenji Nakagawa	4. 巻 E105-B
2. 論文標題 Accurate Parallel Flow Monitoring for Loss Measurements	5. 発行年 2022年
3. 雑誌名 IEICE Transactions on Communications	6. 最初と最後の頁 1530-1539
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transcom.2021EBP3160	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計21件（うち招待講演 2件/うち国際学会 7件）

1. 発表者名 Takahiro Yokoyama, Yoshiki Sato, Sho Tsugawa, Kohei Watabe
2. 発表標題 An Accurate Graph Generative Model with Tunable Features
3. 学会等名 in Proceedings of the 32nd International Conference on Computer Communications and Networks (ICCCN 2023) Poster Session (国際学会)
4. 発表年 2023年

1. 発表者名 Sho Tsugawa, Kohei Watabe
2. 発表標題 Identifying Influential Brokers on Social Media from Social Network Structure
3. 学会等名 in Proceedings of the 17th International AAAI Conference on Web and Social Media (ICWSM 2023) (国際学会)
4. 発表年 2023年

1. 発表者名 Loc Gia Nguyen, Kohei Watabe
2. 発表標題 A Method for Network Intrusion Detection Using Flow Sequence and BERT Framework
3. 学会等名 in Proceedings of 2023 IEEE International Conference on Communication (ICC 2023) (国際学会)
4. 発表年 2023年

1. 発表者名 渡部 康平
2. 発表標題 グラフ生成モデルの発展と今後の展望 ~ 統計的生成モデルからDeep Graph Generatorsまで ~
3. 学会等名 2023年 電子情報通信学会 総合大会 (招待講演)
4. 発表年 2023年

1. 発表者名 横山 昂大, 佐藤 良紀, 津川 翔, 渡部 康平
2. 発表標題 LSTMと条件付きVAEを用いた構造的特徴を高精度に指定可能なグラフ生成モデル
3. 学会等名 電子情報通信学会 情報ネットワーク研究会, 信学技報, vol. 122, no. 342
4. 発表年 2023年

1. 発表者名 ゲン ザ ロック, 渡部 康平
2. 発表標題 フローシーケンスによる侵入検知システムのドメイン適応能力の向上
3. 学会等名 電子情報通信学会 情報ネットワーク研究会, 信学技報, vol. 122, no. 342
4. 発表年 2023年

1. 発表者名 Loc Gia Nguyen, Kohei Watabe
2. 発表標題 Flow-based Network Intrusion Detection Based on BERT Masked Language Model
3. 学会等名 in Proceedings of the 18th International Conference on emerging Networking EXperiments and Technologies (CoNEXT 2022) Student Workshop (国際学会)
4. 発表年 2022年

1. 発表者名 栗山 海渡, 渡部 康平
2. 発表標題 LSTMを用いた任意パラメータ指定可能なネットワークトラフィック生成
3. 学会等名 電子情報通信学会 情報ネットワーク研究会, 信学技報, vol. 122, no. 146
4. 発表年 2022年

1. 発表者名 奥浜 駿, 渡部 康平, 中平 勝也
2. 発表標題 日本におけるマルチビーム衛星通信システムの帯域と電力の配分量について
3. 学会等名 電子情報通信学会 衛星通信研究会, 信学技報, vol. 121, no. 379
4. 発表年 2022年

1. 発表者名 津川 翔, 渡部 康平
2. 発表標題 ソーシャルネットワークにおける情報拡散者の特定に関する一検討
3. 学会等名 電子情報通信学会 コミュニケーションクオリティ研究会, 信学技報, vol. 121, no. 357
4. 発表年 2022年

1. 発表者名 Yuta Yamamoto, Asahi Hentona, Koji Marusaki, Kohei Watabe, Seiya Kawano, Tokimasa Goto, Yutaka Hada, Kazuhisa Fukuzawa, Hirofumi Nonaka
2. 発表標題 Development of the Patent Values Evaluation Method Considering Growth of Technical Community
3. 学会等名 in Proceedings of IEEE Symposium Series on Computational Intelligence (SSCI 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Shohei Nakazawa, Yoshiki Sato, Kenji Nakagawa, Sho Tsugawa, Kohei Watabe
2. 発表標題 A Tunable Model for Graph Generation Using LSTM and Conditional VAE
3. 学会等名 in Proceedings of the 41st IEEE International Conference on Distributed Computing Systems (ICDCS 2021) Poster Track (国際学会)
4. 発表年 2021年

1. 発表者名 荷川取 大, 津波 琉, 渡部 康平, 中平 勝也
2. 発表標題 帯域重畳伝送を適用した無線LAN高効率チャネル割り当て方式の効果
3. 学会等名 2021年 電気学会 全国大会
4. 発表年 2021年

1. 発表者名 中沢 昇平, 佐藤 良紀, 中川 健治, 津川 翔, 渡部 康平
2. 発表標題 LSTMと条件付きVAEを用いた構造的特徴を指定可能なグラフ生成モデル
3. 学会等名 電子情報通信学会 ネットワークシステム研究会, 信学技報, vol. 120, no. 413
4. 発表年 2021年

1. 発表者名 渡部 康平
2. 発表標題 機械学習による通信データ生成へのアプローチ
3. 学会等名 電子情報通信学会 コミュニケーションクオリティ研究会, 信学技報, vol. 120, no. 314 (招待講演)
4. 発表年 2021年

1. 発表者名 中沢昇平, 渡部康平, 中川健治
2. 発表標題 機械学習による実トポロジの特徴を捉えた生成モデルに関する研究
3. 学会等名 2020年 電子情報通信学会 信越支部大会
4. 発表年 2020年

1. 発表者名 栗山海渡, 渡部康平, 中川健治
2. 発表標題 LSTMを用いたGANによる疑似トラヒックの生成に関する - 考察
3. 学会等名 2020年 電子情報通信学会 信越支部大会
4. 発表年 2020年

1. 発表者名 會澤 一輝, 渡部 康平, 中川 健治
2. 発表標題 機械学習を用いた複数のネットワーク間の時系列データの補間に関する研究
3. 学会等名 2020年 電子情報通信学会 信越支部大会
4. 発表年 2020年

1. 発表者名 Kohei Watabe, Masahiro Terauchi, Kenji Nakagawa
2. 発表標題 Model-less Approach for an Accurate Packet Loss Simulation
3. 学会等名 in Proceedings of 2020 IEEE International Conference on Communications (ICC 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 津川翔
2. 発表標題 ノード埋め込みを用いたインフルエンサー推定に関する一考察
3. 学会等名 電子情報通信学会 革新的無線通信技術に関する横断型研究会 (MIKA)
4. 発表年 2020年

1. 発表者名 山本 雄太, 邊土名 朝飛, 中井 堅誠, 作本 猛, 片岡 翔太郎, 野中 尋史
2. 発表標題 技術コミュニティの成長性を加味した特許価値評価手法の開発およびその有効性の検討
3. 学会等名 第83回 情報処理学会 全国大会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	津川 翔 (TSUGAWA Sho) (40632732)	筑波大学・システム情報系・准教授 (12102)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	野中 尋史 (NONAKA Hirofumi) (70544724)	愛知工業大学・経営学部・准教授 (33903)	
研究分担者	中平 勝也 (NAKAHIRA Katsuya) (30500566)	沖縄工業高等専門学校・情報通信システム工学科・准教授 (58001)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関