

# 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

A Study on Pseudo Traffic Data Generation Method with GAN and LSTM

栗山海渡  
Kaito Kuriyama

渡部康平  
Kohei Watabe

中川健治  
Kenji Nakagawa

長岡技術科学大学 電気電子情報工学課程  
Department of Electrical, Electronics and Information Engineering, Nagaoka University of Technology

## 1 はじめに

ネットワークを構築する際，一般的に試験用のトラヒックを生成するトラヒックジェネレータを使用し，ネットワーク機器の負荷テストを事前に行う．しかし，既存のトラヒックジェネレータで実際のトラヒックの特性を多面的に再現した疑似トラヒックを生成することは難しい．そこで，本稿では，機械学習の一種であるGAN(Generative Adversarial Network)とLSTM(Long Short Term Memory)を組み合わせた疑似トラヒック生成手法を提案し，その妥当性を検証する．

## 2 GAN

GANとは，画像などのデータを生成する生成器(Generator)とデータが本物か生成物かを識別する識別器(Discriminator)の二つのニューラルネットワークからなる機械学習の一種である．DiscriminatorとGeneratorが敵対し交互に学習し合うことで，Generatorは本物のデータに類似したデータを生成できるようになる．

## 3 LSTM

LSTMとは，長期的な依存関係を学習できるRNN(Recurrent Neural Network)の一種であり，時系列データに対してよく用いられる．LSTMは，ブロックAがセルCと隠れ層hで鎖状に繋がっている．各ブロックAの内部で，セル状態Cに隠れ層hを追加していくことで，過去の情報を引き継ぐことができる．

## 4 提案手法

本研究で扱うトラヒックデータは，計測期間について一定時間間隔毎に送信されたバイト数の時系列であり，このトラヒックデータを使用し，疑似トラヒックデータを生成する．提案法は，時間間隔毎のバイト数を成分に持つベクトルを入力とし，疑似トラヒックデータを表す同様のベクトルを出力する．

提案手法の構成を図1に示す．まず，LSTMにより，トラヒックデータの時系列的な特性を学習する．次に，GANにより，学習済みのLSTMに入力する最初のセルC，隠れ層h，入力xを学習する．最後に，学習済みのGeneratorによって生成されたセルC'，隠れ層h'，入力x'を学習済みのLSTMの初期値とすることで，疑似トラヒックデータを生成する．

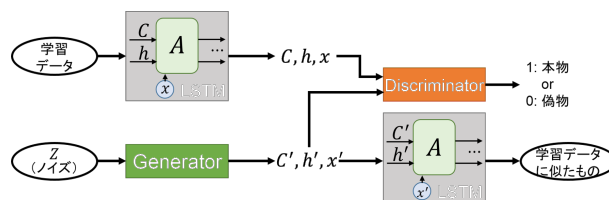
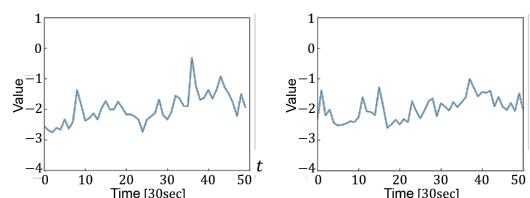


図1 提案手法の構成



(a) 学習データ (b) 生成データ

図2 実験結果

## 5 トラヒックデータ生成実験

公開されているトラヒックデータ[1]を使用し，提案手法を用いて疑似トラヒックデータの生成を行った．トラヒックデータはある1日の30秒毎のバイト数を表しており，前処理として標準化を行った．その後， $t$ を1ずつずらした長さ50のデータを100個生成し，学習データとした．

生成した疑似トラヒックデータを図2に示す．図2より，バイト数が上下する特性を学習できていることが分かる．また，同様の特性を持つ異なるトラヒックデータが生成できていることを確認した．これらのことから，元のトラヒックの特性を有した疑似トラヒックデータが生成できたといえる．

## 6 おわりに

GANとLSTMを用いた疑似トラヒック生成手法を提案し，実際のトラヒックの特性を有した疑似トラヒックデータを生成できた．今後の展望として，具体的な評価手法の検討と，GAN及びLSTMの各パラメータの調整が挙げられる．

## 謝辞

本研究の一部は，JSPS 科研費 20H04172 の助成を受けたものである．

## 参考文献

- [1] Kenjiro Cho *et al.* Traffic data repository at the wide project. In ATEC, 2000.