

コンテンツの需要の時空間パタンの分析と生成

1. 研究背景

先輩の研究

- LEO(低軌道)衛星ネットワークにおける需要の空間的局所性を生かすキャッシュ制御法
 - LEO衛星と、地表のエリアを複数のグループに分類
 - 各衛星は、自身と同じグループのエリアからの要求に対してのみキャッシュを置換
- LEO衛星ネットワーク上で、コンテンツを効率よく広く分散させるためのコンテンツ配置法を実現
 - 各コンテンツのX軸とY軸それぞれの方向で、衛星の配置間隔を遺伝的アルゴリズムを用いて設計
- これらLEO衛星のキャッシュ配信に関する方式の有効性は、地表の各エリアの各コンテンツに対する要求発生パターンに強く依存
 - LEO衛星は通信エリアが限られており、高速で地球を周回しているため、ユーザーとの接続時間が短く、地上の一地点をカバーする時間も限定的
 - こうした制約の中でキャッシュ配信を効率化するためには、いつ・どこで・どのようなコンテンツが必要とされるか(=時空間の需要)を正確に把握することが重要

地域別の需要データが、LEO衛星ネットワークの分析で重要



しかし、これまでに、地域ごとのコンテンツの需要発生パターンに関するデータは測定されていない

2. 研究目的

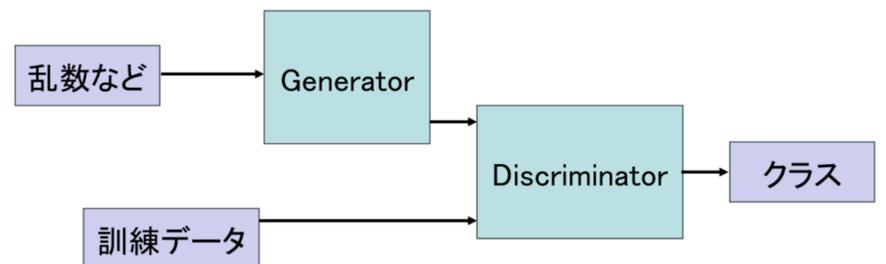
- YouTube Data APIを用いて、国別の高需要コンテンツのパターンや、需要の時系列変化パターンを測定
 - 国別に人気なコンテンツのパターンや、コンテンツの要求数が時間とともにどのように変化するかを測定する
- 時空間パターンを分析
 - 時間と地域を決めたときに、どのような傾向や変化があるのかを分析
- さらに、GANを用いた国別需要データ生成法を提案
 - 機械学習を用いて、需要データを生成

3. GAN

- 敵対的生成ネットワーク(Generative Adversarial Network)
- 2つのニューラルネットワークが競い合うことで、限りなく本物に近いデータを生成する技術

2つのニューラルネットワーク

- 生成器(Generator):
 - 本物のデータと見分けがつかないような偽のデータを生成し、識別器をだますことが目的
- 識別器(Discriminator):
 - 本物のデータと、生成器が生成した偽のデータを正しく見分けることが目的



4. アプローチ

- YouTube Data APIに一定周期でアクセスし、各国の人気動画の上位50個の累積視聴(要求)数を取得
- 差分を取ることで、任意の時間単位の各コンテンツの要求数を取得
- 得られたデータを分析して傾向を把握・可視化
- GANを用いて、国別の需要の時系列データを生成

5. 今後の予定

- YouTube Data APIを用いて、国別に累積要求数が多いコンテンツを測定するためのコードを作成
- プログラムを実行し、累積要求数を測定する
- 取得したデータを分析・可視化
- GANの勉強、コーディングと学習実験