

令和元年6月23日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2015～2018

課題番号：15K21069

研究課題名（和文）曖昧グラフおよびストリーミンググラフにおける問合せに関する研究

研究課題名（英文）Research on queries in uncertain and streaming graphs

研究代表者

佐々木 勇和（Sasaki, Yuya）

大阪大学・情報科学研究科・助教

研究者番号：40745147

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：本研究は曖昧なグラフとストリーミンググラフに対する問合せ技術および関連技術の開発を実施した。曖昧グラフに対する問合せは、グラフの枝に存在確率があるとグラフにおいて、指定した節点が接続している確率を計算するネットワーク信頼性の効率的なアルゴリズムを提案した。また、ストリーミンググラフでは、ソーシャルグラフのように時刻で変化するグラフにおいて、将来のグラフを予測する技術を開発した。さらに、高度な経路探索技術や、確率的なストリーミングデータにおけるパターン検出技術、不完全な道路ネットワークに対する道路の補間技術を提案した。

研究成果の学術的意義や社会的意義

本研究はデータの曖昧性と時系列性をもつグラフの有効な利用方法を開拓し、効率的な計算技術を開発した。曖昧なグラフやストリーミンググラフは通信ネットワーク設計や都市計画など重要な応用例が多い。一方で、計算に掛かる時間やメモリ使用量が大きくなることが多いため、有用にも関わらず曖昧性や時系列変化が有効に活用されていないことが多い。本研究にて提案した効率的な計算アルゴリズムは効果的なグラフの一助となると考えている。

研究成果の概要（英文）：In this research project, we study query processing on uncertain graphs and streaming graphs. In uncertain graphs, we develop an efficient algorithm for computing network reliability, which is a probability that given vertices are connected each other in graphs whose edges have existent probabilities. In streaming graphs, we develop a framework for predicting future graphs. Moreover, we develop an advanced route search query, pattern detection techniques on probabilistic streaming data, and interpolation techniques for incomplete road data.

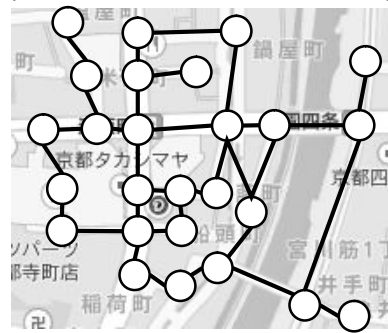
研究分野：情報科学

キーワード：グラフ 曖昧性 時系列 グラフ問合せ

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

世の中の多くのデータがグラフによって表されている。例えば、道路ネットワークや人間関係、セマンティックウェブ、商品流通、分子構造など多岐に渡る。グラフは、学術分野および産業界において非常に重要なデータ構造であるため、国内外問わず、盛んに研究されている。グラフは節点と枝により構成され、枝は節点と節点を繋ぎ、節点間の関連を表す。グラフを構成する節点や枝には、重みやラベルなどの属性値が付くこともある。右図は、京都の一部の道路ネットワークをグラフで表したものである。節点が交差点や建物を表し、枝が交差点や建物間を繋ぐ道路を表す。用途により、節点や枝に付加される属性値が異なる。例えば、節点にはその節点の意味を表すラベルや位置情報など、枝には節点間の距離や移動に必要な時間の重みなどが付加される。



グラフに対する問合せは、日常生活においても一般的になってきている。ナビゲーションシステムを用いた、現在地から目的地までの経路検索は、グラフに対する問合せの典型的なアプリケーションである。他にも、グラフへの問合せでは、到達可能性問合せなど様々なものが提案されている。多種多様な問合せに対して、効率的に計算可能な問合せ技術が長年研究されている。

2. 研究の目的

グラフは曖昧性をもち、時系列的に変化することがある。例えば、通信ネットワークでは、端末間の通信速度（枝の属性値）は一定ではなく、実際の値は送信するまでわからない。また、端末の移動や通信の障害による接続の途切れ（枝がなくなること）や、端末の故障（節点なくなること）も考えられる。さらに、時間経過によりデータが連続的に変化する。これらの特徴を考慮した問合せ技術は、グラフデータの有効活用のためにかかせないため、曖昧性や時系列変化するグラフデータへの問合せ技術を開発する。

曖昧性をもつグラフでは、どのように計算領域を削減するか、およびストリームデータをもつグラフでは、グラフの変化に伴う再計算をどのように削減するかが課題となる。確定的で静的なグラフのための問合せ技術では、解決することができないため、様々な問合せ技術が提案されているが効率性や網羅性の観点で既存研究は十分ではない。

3. 研究の方法

それぞれのグラフ問合せ技術に合わせた計算アルゴリズムを考案し、計算機実験による効率性やメモリ使用量などの定量評価、または問合せの有用性評価のためのユーザ実験を行う。定量評価においては、複数の実世界のグラフデータを利用する。

4. 研究成果

本研究の成果として4つの研究結果を述べる。研究成果として問合せ技術の他に曖昧グラフやストリーミンググラフの応用技術を開発した。

(1) 効率的なネットワーク信頼性計算

ネットワーク信頼性は曖昧グラフにおいて与えられた節点間の接続性を計る指標の一つであり、ネットワーク設計や、都市計画、プロテイン解析など様々な応用がされている。ネットワーク信頼性計算は#P 完全として知られており、既存研究は近似解を計算している。本研究では、サンプリングに基づいた新たなアプローチを提案する。提案アプローチは、層化サンプリングに基づいて精度を落とさずにサンプルサイズを削減し、計算の効率性と高精度をどちらも保つことができる。ネットワーク信頼性の上限値と下限値を用いて、理論的な精度保証を実現する。ネットワーク信頼性の上/下限値を効率的に計算するために、新たな二分決定図であるS2BDDを開発する。S2BDDを構築中に動的計画法を実現し、効率的なサンプリングを実現する。

実験結果より、実データセットを用いて、提案手法が既存のサンプリング手法と比較して最大51.2倍高速（図1参照）であることを確認した。

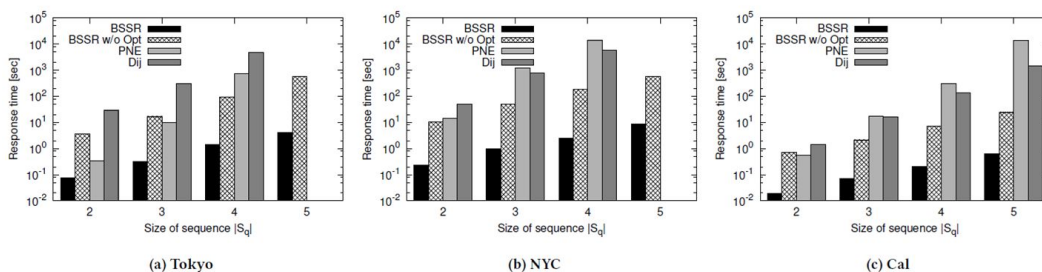


図1：ネットワーク信頼性計算の効率性

(2) スカイラインシーケンスド経路問合せ

旅行計画問合せは、与えられた始点から複数の Point-of-Interest (PoI) を通る経路を探索する問合せである。PoI はユーザ指定したカテゴリに適合する必要がある。旅行計画問合せの既存研究は数多くあるが、ユーザが指定したカテゴリに完全に適合する必要があるため検索の柔軟性が低い。そこで、本研究では、柔軟性が高い旅行計画問合せスカイラインシーケンスド経路 (SkySR) 問合せを提案する。SkySR 問合せは、PoI のカテゴリとユーザ指定のカテゴリの適合度と経路の距離を経路のスコアとして用いて、複数の経路を skyline 問合せの考えをベースとして出力する。

図 2 は SkySR 問合せの主力結果の実例である。青の経路はユーザの指定したカテゴリに完全に適合した最短の経路、赤の経路は指定したカテゴリと十分近く（完全ではない）、青の経路と比べて距離が短い経路である。既存の旅行計画問合せは青の経路しか出力しないが、SkySR 問合せは柔軟な問合せ結果を出力できる。

SkySR 問合せは計算量が大きいため、効率的な計算アルゴリズムを開発した。複数の実データで評価し、最大で 1000 倍以上高速化した。また、SkySR 問合せの有用性を確認するために、ウェブアプリケーションを作成（図 3 参照）し、スペイン・サンタンデルにて実証実験を行い、有用性に関して高評価を得た。

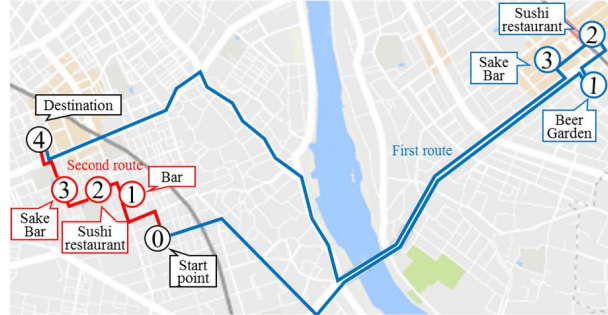


図 2：SkySR 問合せの出力経路例

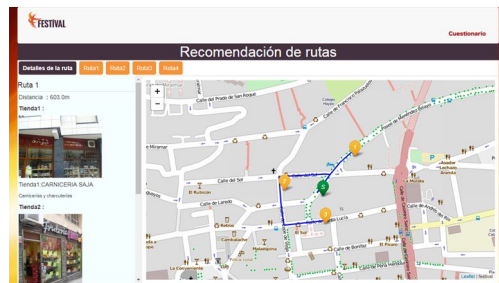


図 3：サンタンデルにおける実証実験用ウェブアプリケーション画面

(3) 不完全な道路データのための道路セグメント補間手法

道路データは位置情報サービスの根幹となる情報である。位置情報サービスを構築する際には、実際の道路と道路データが完全に一致していると考えて設計する。しかし、道路データは、手動更新による遅延や発展途上国における人的リソースの不足などにより実際の道路図と一致していないことが多々ある。そのため、位置情報サービスが有益な結果を提供できない可能性がある。一方で、道路データを自動的に更新する様々な手法が提案されている。本研究では、自動車の移動軌跡データを用いて欠損している道路セグメントを補間する技術を開発する。本技術では、欠損している道路セグメントの候補を確率的なマップマッチング技術により探索し、その後信頼性の高いもののみ補間する。OpenStreetMap とタクシーの移動軌跡データを用いて欠損道路の補間を実施し、実際に欠損道路を補間できていることを確認した。

図 4 は左から補間した道路セグメント、OpenStreetMap の地図、Google map の地図を表す。Google Map の地図では存在しているが OpenStreetMap の地図では存在していない道路セグメントを補間できていることがわかる。

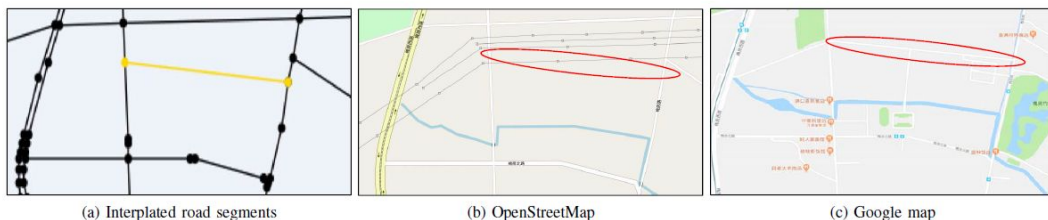


図 4：道路セグメントの補間例

- (4) 深層学習を用いた未知ノードの出現を考慮した時系列グラフの予測
 グラフ構造はソーシャルネットワークなど時間的な変化を伴うことが多い。時系列グラフを対象とした属性予測やリンク予測では、新たに出てくる未知のノードや消失するノードについては考慮していない。本稿では、未知のノードを含む時系列グラフの将来のノード集合、リンク集合、全てのノードが持つ属性を同時に予測する問題について取り組む。そこで、属性予測とリンク予測を時系列グラフ上に拡張した新たな深層学習フレームワークを提案する。提案フレームワークは Graph Neural Network (GNN) と Gated Convolutional Neural Network (Gated CNN) を統合した GNN ブロックを有しており、時系列グラフの属性変化とトポロジ変化を同時に考慮した将来予測を実現する。実世界のデータセットを用いて提案フレームワークによって予測したグラフの精度の検証を行った。

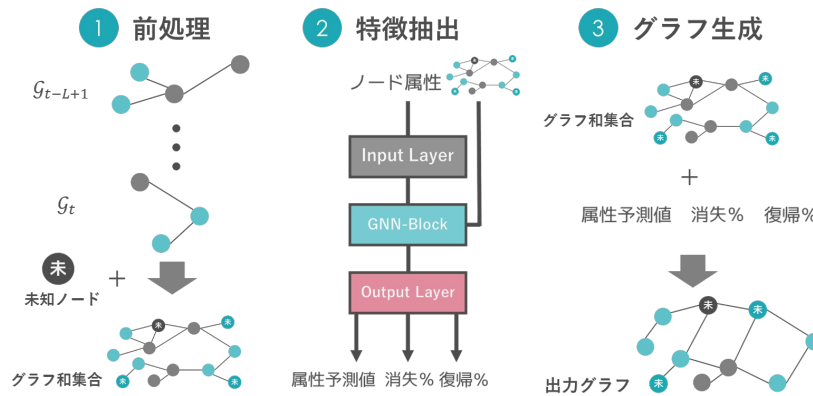


図 4：時系列グラフ予測フレームワーク

5. 主な発表論文等

〔雑誌論文〕(計 1 件)

1. Kento Sugiura, Yoshiharu Ishikawa, and Yuya Sasaki, Grouping Methods for Pattern Matching over Probabilistic Data Streams. IEICE Transactions on Information and Systems, E100-D, 2017, pp.718-729 (査読有)

〔学会発表〕(計 9 件)

1. 山崎翔平, 深層学習を用いた未知ノードの出現を考慮した時系列グラフの予測, 第 11 回データ工学と情報マネジメントに関するフォーラム, 2019
2. 佐々木勇和, Road Segment Interpolation for Incomplete Road Data, IEEE International Conference on Big Data and Smart Computing (BigComp), 2019
3. 佐々木勇和, Efficient Network Reliability Computation in Uncertain Graphs, International Conference on Extending Database Technology (EDBT), 2019
4. 佐々木勇和, Sequenced Route Query with Semantic Hierarchy, International Conference on Extending Database Technology (EDBT), 2018

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

ホームページ等

6. 研究組織

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。