



NTT持株会社ニュースリリース

(報道発表資料)

2013年2月13日

世界最速のグラフデータ分析処理技術を開発 ～日本の人口規模のソーシャルグラフを3分で分析可能に～

日本電信電話株式会社(本社:東京都千代田区、代表取締役社長:鶴浦博夫、以下 NTT)は、ビッグデータ分析の新たな潮流であるグラフデータ分析処理^{※1}において、従来技術を数十倍以上高速化する世界最速の分析技術を開発しました。本技術によりこれまで膨大な時間を必要とした日本の人口規模のソーシャルグラフ^{※2}をわずか3分で分析可能となります。また、世界最大規模の10億ノードのグラフデータを1時間で分析する事も可能となります。これは友人関係を表すソーシャルグラフ・実世界における様々なグラフデータから、最新データを即座に反映したコミュニケーション支援やリコメンドサービスを提供できることを示しています。今後、本技術の適用領域の拡大を通して医療・公共・社会科学分野へ利用シーンの拡大を目指します。

1. 研究の背景

近年、ビジネスにおけるビッグデータの積極的な有効活用がますます重要となっています。その中でもソーシャルメディアやSNSの爆発的な普及に伴い、ビッグデータの持つデータ構造は従来の単純な表構造データから、人・物・場所といった多様な情報のつながり(ウェブ上あるいは実世界のどの場所で利用者が何を参照したか、等)を表現可能なグラフ構造へとシフトしつつあります。このような潮流の変化の中で、グラフデータの持つ様々な情報のつながりを分析することで、個々のデータ間に隠された関係性を捉え、新たなコミュニケーションの支援や個々の利用者にも最適化されたリコメンドへ活用する期待が高まっています。ただし、これまでのグラフデータ分析処理ではグラフデータの大規模性や多様性から分析に膨大な時間を要しており十分な活用ができない課題がありました。

2. 研究の成果

今回、NTTソフトウェアイノベーションセンタでは、大規模なグラフデータを高速に分析可能な「世界最速のグラフデータ分析処理技術」を開発しました。ビッグデータの分析において主要な技術である「クラスタ分析技術」^{※3}および「パーソナライズドページランク技術」^{※4}を用いて、グラフ構造を階層的に集約しながら分析処理を実行することで不要な処理を削減しデータをコンパクト化することで従来技術の数十倍以上の高速化を達成しました。

本技術を活用した事例として、利用者が何をどこで購入した等の購買ログをベースとした商品リコメンドが考えられます。例えば、サッカーに興味を持つ利用者と商品コミュニティにおいて、重要な人物や利用者に適した商品をリコメンドすることができます。従来の技術では、1億ノード規模のグラフデータに対し分析処理に数時間要していましたが、本技術ではわずか3分で分析処理が可能となり大規模なグラフデータを用いた精度の高いリコメンドが行えるようになります。

3. 技術のポイント

(1) 「クラスタ分析技術」(図1)

グラフデータの持つ統計的な情報の変化を動的に捉え、分析しながら計算順序を自動最適化する技術。最適化された計算順序に基づきグラフデータを逐次的に集約し、データ量を削減しながら分析を行う為、本技術を用いるデータの集約パターンは精度劣化が生じないことが証明されている。高い分析精度を維持したまま、計算時間の大幅な短縮が可能となる。

(2) 「パーソナライズドページランク技術」(図2)

ノードの重要度がグラフデータの隣接行列を用いて計算可能であることを利用し、グラフデータのノードを最適に並び替えてから行列分解することでグラフデータをコンパクトな行列へと変換する技術。コンパクトな行列において重要度の低いノードを枝刈りすることにより、重要度の高いノードのみを高速に計算する。

4. 今後の展開

今後は、より大規模なグラフデータへの適用を目指したデータ圧縮技術、より高速なグラフデータ分析の実現に向けた高速化並列化技術などを用いて、本技術の適用領域拡大につとめてまいります。また、本技術の応用例として大規模なソーシャルグラフ・人間の行動履歴を利用することで、新たなソーシャルコミュニケーション支援への応用、商品レコメンデーション、医療・公共・社会科学分野への適用を進めてまいります。

用語解説

※1 グラフデータ分析処理

グラフ構造で表現されたデータに潜む知識や規則性を発見する処理。グラフ構造はデータを表現する手段の1つであり、ノードとノード間の関係を表すリンクから構成される。従来の単純な表構造と異なり、グラフ構造は多様な情報のつながりの表現に適している。

※2 ソーシャルグラフ

人と人との関係をグラフ表現したデータ。ソーシャルグラフでは、ノードは人を表し、ノード間のリンクは友人関係などの人と人との間のつながりを表す。

※3 クラスタ分析技術

グラフデータの持つコミュニティ構造に従って、互いに類似した特徴を有するクラスタと呼ばれる部分にグラフデータを分割する技術。従来の最速技術として、ルーヴァン・カトリック大学、ピエール・エ・マリー・キュリー大学、インペリアル・カレッジ・ロンドンのチームが開発したアルゴリズムが挙げられる。

※4 パーソナライズドページランク技術

グラフデータの一部のノード群を指定して重みづけをすることで、グラフデータの持つデータ間の関係構造に従ってグラフデータを分析し、指定したノード群からの影響の高いノード群を決定する技術。従来の最速技術として、インド工科大が開発したBasic Push Algorithm や、INRIA と twente大のチームが開発したMonte Carlo法が挙げられる。

別紙・参考資料

- ▶ [図1 クラスタ分析技術のポイント](#)
- ▶ [図2 パーソナライズドページランク技術のポイント](#)

本件に関するお問い合わせ先

■ 日本電信電話株式会社
サービスイノベーション総合研究所
企画部広報担当
E-mail: randd@lab.ntt.co.jp
Tel: 046-859-2032

ニュースリリースに記載している情報は、発表日時点のものです。現時点では、発表日時点での情報と異なる場合がありますので、あらかじめご了承くださいとともに、ご注意をお願いいたします。

NTT持株会社
ニュースリリース

▶ [最新ニュースリリース](#)

▶ [バックナンバー](#)

▶ [English is Here](#)

NTT持株会社
ニュースリリース内検索

1997 ▼ 年 04 ▼

月 ~

2021 ▼ 年 11 ▼ 月

検索

NTTグループの情報は
こちらからもご覧いただけます。



▲ [このページの先頭へ](#)

▶ [更新履歴](#) ▶ [サイトマップ](#) ▶ [お問い合わせ](#) ▶ [著作権](#) ▶ [プライバシーポリシー](#) ▶ [情報セキュリティポリシー](#) ▶ [ウェブアクセシビリティポリシー](#) ▶ [個人情報保護について](#)

Copyright © 2021 日本電信電話株式会社