DEIM Forum 2022 G23-4 (day2 p37)

知識グラフにおける更新エンティティの予測

大倉真一希† 天笠 俊之††

† 筑波大学大学院システム情報工学研究科 〒 305-8577 茨城県つくば市天王台 1—1—1
†† 筑波大学計算科学研究センター 〒 305-8577 茨城県つくば市天王台 1—1—1
E-mail: †okura@kde.cs.tsukuba.ac.jp, ††amagasa@cs.tsukuba.ac.jp

あらまし 近年，知識グラフへの注目度が高まっており，様々な分野において応用されている．一方で知識グラフは，本質的に不完全であり継続的なメンテナンスが不可欠であるという性質も抱えている．このメンテナンスの自動化のため様々な研究が行われているが，依然として人手に依存する部分は大きい．そこで本研究では，自動化手法及び人手でのメンテナンスに共通する問題点である，膨大なエンティティを含む知識グラフ中から，更新を行うべきものを発見するのが難しいという点に着目した．知識グラフの構造情報や直近の更新情報をもとに新たに更新を行うべきエンティティを予測できれば，更新すべき箇所を探す作業を効率化できるとともに，更新忘れなどの対策も可能となる．現在までに汎用知識グラフである Wikidata を元とした編集履歴を含むデータセットの構築及び，リンク関係や編集履歴を元に更新が起きるエンティティを予測する手法の開発を行い，実験によってそれらの有用性を確認した．

キーワード 知識グラフ，更新予測，Personalized PageRank

1 はじめに

近年，知識グラフへの注目度が高まっている．知識グラフとは，世の中のエンティティ同士の関係を主語 (subject, head entity)，述語 (predicate, relation)，目的語 (object, tail) から成るトリプルの形式によって表したグラフ構造である．知識グラフの応用例は広く，例えば自然言語処理の分野では質問応答（QA）タスクの回答データベースや機械学習タスクの教師データとして，画像認識の分野では画像中の物体同士の関係理解，推薦システムの分野ではコールドスタート問題の対策や推薦結果への情報補完などにそれぞれ用いられている．

知識グラフは，上述したような利点を持つ一方，本質的に不完全であり，継続的なメンテナンスが不可欠という性質も抱えている．メンテナンスが必要な例としては，現実世界の事実情報に追随したリレーションの追加・削除・更新，新規エンティティの追加，不備によりリレーションが欠けてしまっているトリプルの補完などが挙げられる．このような問題へ対処するために様々な研究が行われている．リレーションの更新や，新規エンティティの追加に関する研究としては，Web 文書などのパブリックかつ非構造なデータから新たなトリプルの抽出を行う OpenIE [1] や，抽出したトリプルを知識グラフ中に取り込むために，エンティティの同一性の判定を行う Entity Recognition [2] などがある．また，欠けたトリプルの補完としては，トリプルのうちの 2 要素から，欠けた 1 要素の予測を行う，知識グラフ補完に関する研究が盛んに行われている [3]?[5].

このように，知識グラフのメンテナンスを自動で行うことを目的として，様々な分野の研究が行われている．しかし一方で，現状の知識グラフのメンテナンスは，依然として人間の手に大きく依存しており，前述した自動化のための手法も，多くは人手と組み合わせて用いられる．

人手と自動化手法によるメンテナンスの両方に共通する問題点として，膨大なエンティティを含む知識グラフの中から，メンテナンスを行うべき箇所を判定するのが難しいという点が考えられる．人手でのメンテナンス行う場合，対象のエンティティの特定は，各人が個人で得た情報に依存する．また，自動化手法を用いて非構造文書などよりトリプルを追加する場合も，追加する情報はソースのドメインに依存する．よっていずれの手法を用いる場合においても，知識グラフ中の更新を行うべき箇所を網羅できていないと言える．もし，知識グラフ中の更新すべきエンティティを自動で検出できれば，人手によるメンテナンスにおける個人の知識差による影響を減らすことができ，また自動化手法で見落とされた箇所などの補完も可能となる．

よって本研究では，知識グラフにおいて，情報を更新すべきエンティティを自動で検出する手法の開発を目的とする．今回は，あるエンティティへの編集が，そのエンティティと周辺エンティティへの更なる編集を誘発すると考え，エンティティのリンク関係及びそれぞれへの編集情報を元にした手法の提案を行った．また，提案手法の有効性を確認するための評価実験を行い，エンティティのリンク関係を用いたベースライン手法との性能を比較検討した．

2 関連研究

2.1 知識グラフの状態遷移判定

Wijaya ら [6] は，知識グラフ中のエンティティについて，それが新たな事実情報の追加などが起きる状態にあるかを Contextual Temporal Profiles（CTPs）を用いて判定する手法の提案を行った．CTPs とは，各年代ごとにおけるエンティティの特性を，その年代に現れた文書からなるコーパスから推定したようなものである．Wijaya らは，CTPs の設計のためにまず，各年代のコーパスから，事象の変化の元となるようなエン

ティティ（大統領など）の周辺に現れる語句を抽出し，そのエンティティのドメインの特性を表現するベクトルの生成を行った．次に，年代ごとに得られたドメインの特性ベクトルの差分を考えることで，ドメインの遷移状態を表すようなベクトルを得た．最後に，ドメインの新規エンティティについて，その特性ベクトルなどを生成し，得られたドメインの特性ベクトル及び，遷移状態ベクトルと比較することで，そのエンティティに関する知識トリプルが発生する年代の推定を行った．この研究では，知識グラフのエンティティの遷移状態を大規模コーパスを元にモデル化し，変化を予測することを可能としたが，ドメインを表すような seed エンティティの選定を人手で選択する必要がある点や，モデル化のためのソースが年代ごとのコーパスという長期的なものに依存しているなどの点で改善の余地がある．

2.2 知識グラフにおけるリンク予測

知識グラフ中のエンティティについてリンク予測を行うタスクは知識グラフ補完（Knowledge Graph Completion）と呼ばれ，盛んに研究が行われている．知識グラフ補完の代表的な手法として，Bordes らが提案した TransE [3] がある．TransE は，知識グラフ中のトリプル (h, r, t) をもとに，エンティティに対するベクトル $\boldsymbol{e}_i$，リレーションに対するベクトル $\boldsymbol{l}_j$ を，$\boldsymbol{e}_h + \boldsymbol{l}_r = \boldsymbol{e}_t$ の関係を満たすように学習する手法である．この手法を用いて，$(h, r, ?)$ のような一部の要素が欠けたトリプルについて，ベクトル $\boldsymbol{e}_h + \boldsymbol{l}_r$ と相関が高いベクトル $\boldsymbol{e}_t$ を求めることで，tail になり得るエンティティの候補を得ることができる．

TransE は精度面，応用面で様々な拡張がされている [4], [5]．その中でも Leblay ら [7] は，時間情報を考慮した知識グラフ補完のタスクに取り組んでいる．このタスクにおいては，通常のトリプルに時間情報 τ を追加したものを扱う．時間情報も考慮した埋め込み表現を求めることで，トリプルの発生する時間の予測 (（Ethiopia, Praise or endorse, China, ?）) や，ある時間が指定された場合のリンクの予測 (（Canada, Host a visit, ?, 2014-04-20）) が可能となる．

このように知識グラフにおけるリンク予測に関して様々な研究が行われている．しかし，いずれにおいても，トリプルのうちの一部が指定されることを前提としており，予測を行うべきエンティティの検出については考えられていない．

3 提 案 手 法

3.1 手法の方針

手法の方針として，まず，知識グラフにおいて，どのようなエンティティに編集が行われるかについて考える．ある時刻 t とそこから α だけ経過した時刻 $t+\alpha$ を比較した時，更新が行われたエンティティ及びその周辺にあるエンティティは今後においても更新がされやすいと想定される．例として，知識グラフ中のエンティティ *Joe Binden* に新たな関係，(*Joe Binden, position hold, President of U.S*) が追加されたケースを考える．この時，新たな関係が追加された *Joe Biden* は，大統領就任に伴い周辺に様々な変化が予想されるため，知識ベース上でも今後更新がされやすいと考えられる．また，その周辺にあるエンティティにも変化が伝播していくことが予想される．更に，編集の伝播は，時間の経過とともに減衰していくと考えられる．

次に，エンティティの編集を伝播しやすいようなリレーションについて考える．ある時刻 t においてエンティティに編集が行われた時，その編集は，過去により多くの編集を伝播したリレーションによって伝播されると考えられる．例えば，知識グラフ知識グラフ中のエンティティ *President of U.S* に新たな編集が加えられると，それに続いて自身やその周辺の詳細情報、親族などにさらなる編集が加えられると予想される。一方、出身地や言語など、エンティティに依存しない関係は編集の伝播を誘発しない。このような，過去に編集を伝播したかの履歴をもとに，リレーションへの重みを決定する．

また，編集の伝播は，より多くに参照される上位概念のエンティティから，参照の少ない下位概念のエンティティの方向に発生し，その逆はほとんどないことが予想される．例として，知識グラフ中のエンティティ *Joe Binden* に編集が起こったケースについて考える（図 1．このとき，エンティティ *Joe Binden* と同じ人間であり，被参照数が比較的少ないエンティティ *Jean Binden* には編集が伝播することが予想されるが，上位概念で，被参照数も非常に多いエンティティ *human* に編集が伝播することはないと考えられる．この性質をもとに，編集履歴をもとにしたリレーションへの重みに更なる重みづけを行う．

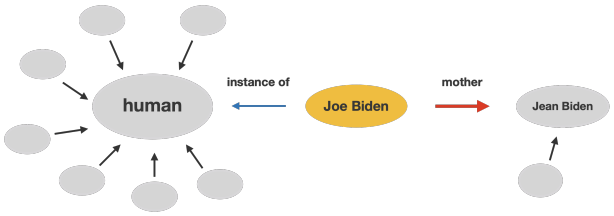

図 1 エンティティ間のリレーションの重み付けに関する基本アイデア．編集の伝播は，より多くから参照される大きい概念のエンティティから参照数の少ない小さい概念のエンティティの向きへ起こると考えられるため，それを考慮した重み付けを行う．

3.2 手法の設計

これらの仮定を元に手法の設計を行う．今回は，前述した仮定を取り入れつつ，編集が起きたエンティティに類似したエンティティを見つけ出すために，Personalized PageRank（PPR）[8] を手法のベースとして用いることとした．PPR は，クエリノードを始点とするランダムウォークを行い，その際，一定確率 α でクエリノードのいずれかにその問い合わせ確率に応じてジャンプすることで，グラフ中のノード上でのランダムサーファーの存在確率を算出し，それをクエリノードとの類似度として用いる手法である．

グラフのノード数を N，要素 i がノードの問い合わせ確率を表すクエリベクトルを $\boldsymbol{b} \in \mathbb{N}$，各列が，ノードの隣接ノードへの遷移確率を表すベクトルである隣接行列を $\boldsymbol{A}$，最終的に得ら

れるランダムサーファーの各ノード上での存在確率のベクトルを $\boldsymbol{v} \in \mathbb{N}$ とすると，定常状態における $\boldsymbol{v}$ は，以下の式を収束するまで再起的に計算することで得られる．

$$\boldsymbol{v} = (1-\alpha)\boldsymbol{A}\boldsymbol{v} + \alpha\boldsymbol{b} \tag{1}$$

提案手法においては，前述した仮定を PPR に取り入れつつ，より推測がしづらい（定常的でない）エンティティへの編集を捉えられるよう，エンティティ e_i への編集の非定常性のスコアを定義し，それをもとに隣接行列 $\boldsymbol{A}$ 及びクエリノード $\boldsymbol{b}$ にそれぞれ次のような変更，制約を加えることとする．

3.2.1 エンティティへの編集の非定常性スコア

知識グラフ中のエンティティへの編集の中でも，定常的に行われているものではなく，イベント的に発生するものを捉えられるよう，エンティティ e_i に対する編集の非定常さを表現するスコア $escore(e_i)$ を設計する．今回は次の要素をもとにスコアを決定することとした．

（1） **前期間と現期間の編集数の比率**．エンティティに非定常な編集が起こっているかは，前期間と現期間の被編集数の比率に反映されていると考えられる．一方，直近 1 週間と 2 週間前のように，日時が近くかつ短い期間同士を比較した場合，両方の期間に同じ性質の編集が含まれ，非定常性が比率に反映されないことが起こり得る．よって今回は，現期間を直近 1 週間，前期間を基準日の前月とし，比率の算出を行うこととした．

（2） **エンティティへのリンク数**．リンク数が多いエンティティの場合，編集の対象が多いことから，編集数もそれに応じて増加すると考えられる．よって，リンク数を用いて前述したエンティティへの編集数の比率を標準化することとした．

上記をもとに，(2) 式のようにエンティティへの編集の非定常性スコアを定義した．ここで，$N_{out}(e_i)$ はエンティティ e_i の外向きリンク数，$D_w(e_i)$ は基準日の直近 7 日間におけるエンティティ e_i の編集回数，$D_{lm}(e_i)$ は基準日の前月におけるエンティティ e_i の編集回数である．

$$escore(e_i) = \frac{D_w(e_i)}{N_{out}(e_i) D_{lm}(e_i)} \tag{2}$$

3.2.2 隣接行列 $\boldsymbol{A}$

知識グラフ中のリレーションに対して，関連するエンティティが編集されるケースの総数から，リレーション r_j のスコア $rscore(r_j)$ を (3) 式のように計算する．ここで E_{edited} は予測日の 1 週間前までに編集が起きたエンティティの集合，$E_{(e_i,r_j)}$ はエンティティ e_i にリレーション r_j でリンクされており，かつ予測日の 1 週間前〜2 週間前の間に編集が起きたエンティティの集合，$N_{in}(e_i)$ はエンティティ e_i の内向きリンク数である．

$$rscore(r_j) = \frac{1}{|E_{(e_h,r_j)}|} \sum_{e_i \in E_{edited}} \frac{N_{in}(e_i)}{N_{in}(e_j)} \sum_{e_k \in E_{(e_i,r_j)}} escore(e_k) \tag{3}$$

$rscore$ をもとに，エンティティ e_i, e_j 間の重み $W_{(e_i,e_j)}$ を (5) 式のように定義する．ここで，E は全エンティティの集合，E_{e_i} はエンティティ e_i とリンクされているエンティティの集合，R は全リレーションの集合，$R_{(e_i,e_j)}$ はエンティティ e_i, e_j 間を繋ぐリレーションの集合である．

$$s(r_i) = rscore(r_i) + \min\{rscore(r) \mid \exists r \in R, rscore(r) != 0\} \tag{4}$$

$$W_{(e_i,e_j)} = \frac{\frac{N_{in}(e_i)}{N_{in}(e_j)} \sum_{r_l \in R_{(e_i,e_j)}} s(r_l)}{\sum_{e_k \in E} \frac{N_{in}(e_i)}{N_{in}(e_k)} \sum_{r_m \in R_{(e_i,e_k)}} s(r_m)} \tag{5}$$

3.2.3 クエリベクトル $\boldsymbol{b}$

知識グラフ中の編集が起きたエンティティについて，行われた編集の数が多いかつ，直近で編集が行われたものほど，高い問い合わせ確率を与えるように設計を行う．エンティティ e_i に対応するクエリベクトル $\boldsymbol{b}$ の要素を b_{e_i}，期間内でのエンティティ e_i の編集履歴の集合を H_{e_i}，編集 d の発生時刻（時間）を $t(d)$，基準時刻（時間）を t_{base} とすると，(6) 式より表される．

$$b_{e_i} = \sum_{d \in D_{e_i}} \frac{1}{t_{base} - t(d) + 1} \tag{6}$$

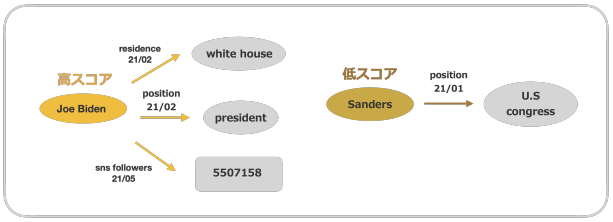

図 2　クエリベクトル設計のための基本アイデア．より直近に，多くの編集が行われたエンティティに高い値を与える．

4 実　　験

4.1 比較手法

4.1.1 エンティティに付随するリンクの数を用いる手法

エンティティは，それに付随するリンク数が多いほど注目が大きくまた，更新箇所も多くなるため，更新頻度が高くなると考えられる．よって，エンティティのリンク数をその更新されやすさを表す値とみなし，手法の一つとして用いることとした．本手法では，エンティティのリンク数が閾値を超えたものを更新が起きると見なす．リンク数としては，エンティティの外向きのリンク数 (nlink-out)，内向きのリンク数 (nlink-in)，内向き・外向きのリンク数の和 (nlink-all) の 3 つのパターンを用いることとした．

図 3 に例を示す．エンティティ *President* には，内向きのリンクとして *position* が 2 つ，外向きのリンクとして *replace* が 1 つあるため，$in = 2, out = 1, all = 3$ となる．

4.1.2 エンティティの直近の編集数を用いる手法

直近で多くの編集が行われたエンティティは，今後においても編集が起きやすいと考えられる．よって，エンティティの過去の編集数を更新のされやすさとみなし，それが閾値以上となるものを更新が起きると予測する手法 (nedit) を比較手法の一つとして採用した．

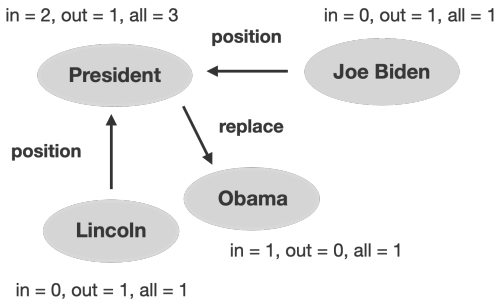

図 3　エンティティのリンク数の集計例.

4.1.3　PageRank を用いる手法

知識グラフ中のエンティティは，より多くのエンティティに参照されるようなエンティティほど，重要度が高く，更新がされやすいと考えられる．そのような考え方をもとに，グラフ中のノードのスコアを算出するアルゴリズムとして，PargeRank [9] がある．PageRank は，多くのノードにリンクされているノードからリンクされているノードほど重要度が高いというアイデアを基に，グラフ中のノードのスコアを計算するアルゴリズムである．グラフ中のノード A に対するページランク値 $PR(A)$ は，ノード A にリンクされているノードを $T_i(1 \sim n)$，ノード Ti にリンクされているノードの総数を $C(T_i)$，ハイパラーパラメータを d とすると，以下の式より表される．

$$PR(A) = (1-d) + d\sum_{i=1}^{n} \frac{PR(T_i)}{C(T_i)} \qquad (7)$$

今回は，知識グラフを有向グラフと見なして，各エンティティの PageRank スコアを計算し，スコアが閾値以上となったエンティティを更新が起きると予測することとする．

4.2　提案手法のうちクエリベクトル，隣接行列のみを用いる手法

提案手法におけるクエリベクトル及び隣接行列それぞれの効果を確認するため，隣接行列のみを利用する手法（proposed-w），クエリベクトルのみを利用する手法（proposed-qv）を比較手法の 1 つとして採用する．

4.3　データセット

実験用データセットとしては，手法の有用性の検証に重要となる要素である，1. 知識グラフ全体の時系列での変化を見れる形でデータが公開されている，2. 各エンティティの更新履歴をその操作時間と共に取得することができる，などを満たしているかを考慮し，公的に公開されている共同編集型大規模知識グラフの一つである Wikidata [10] をもとに作成することとした [1]．Wikidata では，グラフ全体の dump データをおよそ 1 週間ごとに提供しており [2]，また，各エンティティの編集履歴を，それぞれの詳細ページより取得することができる．

実際のデータセットの構築は以下の手順で行なった．

1：https://www.wikidata.org

2：https://dumps.wikimedia.org/wikidatawiki/entities/

（1） Wikidata dump のダウンロードページより，2021/05/24，2021/05/31，それぞれの dump を取得する．

（2） 2021/05/24 中のグラフより，データ中の先頭 100 個の head エンティティを取り出す．

（3） 得られた 100 個のエンティティを起点として，4 ホップ先までに出てきたエンティティと，それを head として持つトリプルをデータセットとして利用する．

（4） 最終的に得られた全てのエンティティの編集履歴を各エンティティの詳細ページよりスクレイピングし取得する．

この手順の結果として得られたデータセットの内訳は表 1 の通りである．

表 1　Wikidata より構築したデータセットの内訳

	2021/05/24	2021/05/31
トリプル数	8,487,493	8,495,281
編集が起きたエンティティ数	62751 (5/25〜5/31)	48379 (6/1〜6/7)
エンティティ総数	995,012	
リレーション総数	5,782	

4.4　更新エンティティの予測に関する実験

知識グラフ中のエンティティ数は膨大であるため，そこから更新が起こり得るものを提示することは，実際のメンテナンスにおいても有用と考えられる．よって，前述した手法がどの程度の精度で，知識グラフ中のエンティティへの更新を予測できるかの実験を行った．

4.5　評価指標

評価指標としては以下を用いる．

- 編集されると予測したエンティティ e_i の $escore(e_i)$ の平均値.
- 手法によって算出したエンティティの重要度が上位 100 件のうち，実際に編集が行われたものの割合.
- 手法によって算出した重要度が閾値を超えたエンティティを編集が起きると予測する.
  - 予測結果において，test データ中で更新が行われたエンティティのうち何割を検出できているかを Recall として評価する.
  - 更新が起きると予測されたエンティティのうち，実際に何割に更新が起きたかを Precision として評価する.

4.6　実験手順

実験は次の手順で行った.

（1） 2021/05/24 のグラフに対して，2021/05/25 〜 2021/05/31 の間で編集が起きるかを予測し，それをもとに *escore* の平均値，Recall，Precision それぞれが最大に成るように，ハイパーパラメータ最適化フレームワークである Optuna [11] の多目的最適化機能を用いて，閾値のパレート解を得る（過去の編集履歴として，nedit では 2021/05/17 〜 2021/05/24 のものを，提案手法では 2021/05/10 〜 2021/05/17 と 2021/05/17 〜 2021/05/24 のものを用いる）.

（2） 2021/05/31 のグラフに対して，得られた閾値のうち，平均 score を元に，2021/06/01 〜 2021/06/07 の間で編集が行われるエンティティを予測し，性能を評価する（過去の編集履歴として，nedit では 2021/05/24 〜 2021/05/31 のものを，提案手法では 2021/05/17 〜 2021/05/24 と 2021/05/24 〜 2021/05/31 のものを用いる）．

4.7 実験結果

上記の設定で行った実験の結果を表 2，3 に示す．

表 2 実験結果（Recall が最大となる閾値の場合）

	Top 100	Precision	Recall	平均 escore*
nlink-in	38 個	29.4%	0.467%	0.19
nlink-out	38 個	15.9%	**8.13%**	0.0868
nlink-all	38 個	27.7%	0.961%	0.185
nedit	43 個	17.4%	2.15%	0.173
pagerank	**49 個**	**47.3%**	0.104%	0.224
proposed-w	21 個	8.62%	0.885%	0.0833
proposed-qv	39 個	42.9%	0.105%	**4.98**
proposed	25 個	30.1%	0.875%	2.08

* 全編集を正確に予測した際の平均 escore は 1.24 となる．

表 3 実験結果（平均 escore が最大となる閾値の場合）

	Top 100	Precision	Recall	平均 escore*
nlink-in	38 個	66.7%	0.0239%	0.72
nlink-out	38 個	31.8%	**0.653%**	0.259
nlink-all	38 個	66.7%	0.0239%	0.72
nedit	43 個	47.8%	0.0145%	1.35
pagerank	**49 個**	64.3%	0.0153%	0.721
proposed-w	21 個	16.7%	0.00251%	0.333
proposed-qv	39 個	**70.6%**	0.0301%	8.95
proposed	25 個	36.4%	0.01%	**15.0**

* 全編集を正確に予測した際の平均 escore は 1.24 となる．

まず，表 2 の Recall が最大となるパレート解の閾値を用いた際の結果を見ると，Recall は外向きリンク数を用いる手法にて最も高くなっており，平均 escore は提案手法にて最大となっていることがわかる．一方，表 3 をもとに，平均 escore に注目して比較すると，提案手法が圧倒的に良い結果となっている．これより，検出精度の観点では，更新エンティティの検出数を重視するならば外向きリンクを用いる手法，非定常的なエンティティへの更新の検出を重視するならば，提案手法を用いるのが良いと言える．

また，提案手法と proposed-w，proposed-qv を比較すると，提案手法は proposed-qv よりも Precision などの観点で良いが，Recall の観点で悪く，proposed-w よりはほとんどの観点で優っていることがわかる．このような結果となった原因としては，proposed-qv はより直近で多くの編集が起きたものを編集が起きると予測するため編集トレンドを捉えやすく精度が高くなる一方，隣接行列のみを利用する場合は，リレーションについてしか考慮していないため，予測前期間で編集が起きていないかつ，次数が大きい上位概念のエンティティのリンク先のマイナーなエンティティを編集が起きると予測していることが考えられる．しかし，proposed-qv に proposed-w を組み合わせた提案手法では Recall が上昇していることから，単純に編集が行われたものを高スコアにするだけでは予測できない編集を捉えられていると言える．

次に，各手法における検出数に対する精度と検出した編集の非定常性の変化を見ることとした．Recall の変化に対する Precision の遷移を結果を図 4 に，Recall の変化に対する平均 escore の遷移を結果を図 5 にそれぞれ示す．結果を見ると，Recall が小さい領域において，Precision に関してはリンク数手法が優っており，平均 escore に関しては提案手法が良いという差が見られる．しかし，全体を見ると大きな違いにはなっていない．また，PageRank 及び提案手法は Recall が中間程度となる閾値が存在せず，多くのエンティティを検出するのが難しいと言える．このような結果となった原因としては，Wikidata に対して行われる編集の多くがグラフ上の情報だけでは予測が難しいものであり，Recall が一定以上となる領域においては，予測の成否がほぼランダムになってしまうためと考えられる．

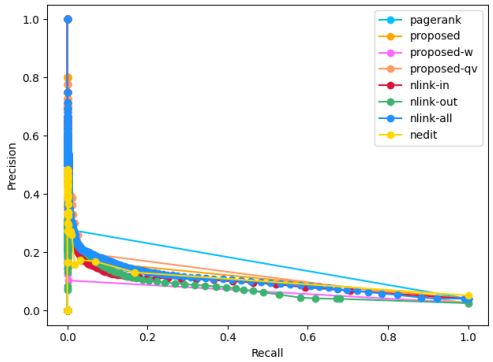

図 4 各手法における Precision と Recall の関係

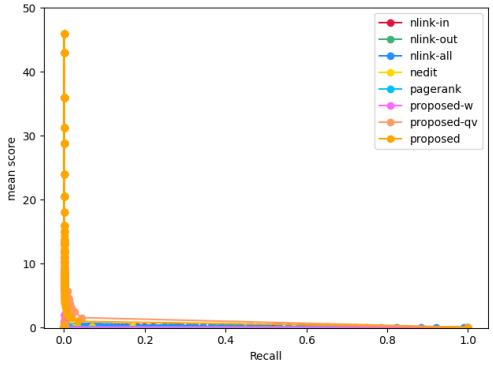

図 5 各手法における平均 escore と Recall の関係

4.8 各手法の予測結果の解析

4.8.1 各手法において高いスコア付をされたエンティティの例

予測結果の解析としてまず，各手法の予測結果の傾向を見る

ために，どのようなエンティティが高いスコア付をされているかを確認することとした．表 4 に，各手法においてスコアがトップ 5 となったエンティティの一覧を示す．

表 4 を見ると，nlink，pagerank，proposed-w，proposed-qv は似通った結果になっており，参照数の多い上位概念のエンティティが高いスコアとなっていることがわかる．一方，nedit は，航空機に関するエンティティが多く提示されており，編集トレンドが反映されていると言える．最後に，提案手法のスコアは直近の編集履歴の影響が大きく，その上で参照数などが考慮されている傾向が見て取れた．これより，知識グラフ内で影響力の大きいエンティティを見つけるならば，nlink，pagerank を，直近の編集トレンドをもとにエンティティを提示するならば nedit を，その両方を考慮するならば提案手法を用いるのが良いと考えられる．

表 4　各手法において高いスコア付をされたエンティティ

nlink-in	**nlink-out**	**nlink-all**	**nedit**
human	The Complete Encyclopedia of World Aircraft	human	Pratt & Whitney F100
Wikimedia category	Sol	Wikimedia category	McDonnell Douglas F-15 Eagle
People's Republic of China	2017 German presidential election	People's Republic of China	McDonnell Douglas F-15E Strike Eagle
UTC+08:00	1999 German presidential election	UTC+08:00	AN/APQ-181
male	2010 German presidential election	male	AN/APQ-153
pagerank	**proposed-w**	**proposed-qv**	**proposed**
Wikimedia category	UTC+08:00	Wikimedia category	The Cage
human	Wikimedia list article	human	Igbo people
People's Republic of China	mayor of a place in France	People's Republic of China	UTC+08:00
UTC+08:00	Wikimedia category	UTC+08:00	Maryland Center for History and Culture
village-level division in China	family name	village-level division in China	Wikimedia list article

4.8.2　各手法によって検出された更新エンティティの重複数

現在までの解析より，各手法によって検出されたエンティティには，性質の違いがあることがわかった．そのため，検出結果に重複が少ない手法を併用することにより，より多くの更新エンティティの検出ができると考えられる．各手法の併用の有効性を確認するために，検出したエンティティがどれだけ重複していないかを検証した．表 5 にその結果を示す．ここで，表中の値は，左列のラベルが示す手法が検出したエンティティに対して，各列のラベルの手法で検出できていなかった数を表している．結果を見てみると，nlink-out，nedit，proposed-w，提案手法によって検出されたエンティティは，他の手法によって検出されたものとの重複が少ないことがわかる．特に nedit，提案手法では他手法との重複が 15%以下となっている．よって，その他の手法との併用が最も有効であると言える．

また，proposed-qv の結果に注目すると，リンク数手法，pagerank との重複が多いことから，より編集が行われ辛いマイナーなエンティティを推薦するためには proposed-w と組み合わせるのが有効であると言える．

4.9　編集伝播の予測

知識グラフ中のあるエンティティの更新が行われた時，その周辺の関連度が高く，編集可能性が高いエンティティを推薦するのは，実際のシステムにおいても有用だと考えられる．そこで，各手法を用いて更新が起きたエンティティの周辺エンティティのスコア順で並べたとき，有用な推薦結果が得られるか実験を行った．

4.9.1　実 験 手 順

実験は次の手順で行った．

（1） 2021/05/24 のグラフ及び，2021/05/25 〜 2021/05/31 の編集履歴を用いて，各手法のスコアを算出する．

（2） 2021/05/31 のグラフ中で，2021/06/01 〜 2021/06/07 の間で編集が行われたエンティティである Q6279 (Joe Biden)，Q1490 (Tokyo) について，その周辺エンティティをスコア順に出力し，その有用性を確認する．

4.9.2　実 験 結 果

各手法によるエンティティ Q6279 (Joe Biden) の周辺エンティティのランク付の結果を表 6 に，Q1490 (Tokyo) に対する結果を表 7 ににそれぞれ示す．

まず，個人に関するエンティティ Q6279 (Joe Biden) に対する結果である表 6 を見ると，nlink, pagerank では，居住地や政党など，個人よりも概念が上位のエンティティが高いランクとなっている傾向が見てとれ，nedit，提案手法では同程度の階層の概念である個人，親族に関するエンティティが推薦されていることがわかる．特に，提案手法では，より編集の伝播がされやすいことが予想される *child*, *sibling* に高いランク付がされていることから，リレーションごとの編集の伝播のしやすさも考慮できていると考えられる．また，proposed-w，proposed-qv，提案手法を比較してみると，提案手法の結果が proposed-qv の結果に似通っており，隣接行列よりもクエリベクトルに大きく影響を受けていることが推察される．

次に都市に関するエンティティ Q1490 (Tokyo) に対する結果である表 7 に注目する．結果を見ると，nlink-out, nlink-all では都知事など，nlink-in, nedit では行政区域などが高いランクとなっていることがわかる．提案手法では付随して編集が行われる可能性が高いと思われる本部や，行政区域，年表から直近で編集が行われたものが高ランクになっている．どの手法の推薦結果も，元エンティティの編集に関連する物としては妥当であるが．情報を増やすという観点では，よりマイナーなものを推薦している提案手法が有用と考えられる．

5　まとめと今後の課題

本研究では，知識グラフにおける新たなリンクの追加や属性値の修正などの編集が行われるエンティティを検出することを

表 5　各手法によって検出されたエンティティについて，他の手法で検出したものとの非重複数の一覧.

比較元＼比較先	nlink-in	nlink-out	nlink-all	nedit	pagerank	proposed-w	proposed-qv	proposed
nlink-in	0 (0.0%)	47 (58.8%)	1 (1.25%)	80 (100%)	17 (21.2%)	56 (70.0%)	24 (30.0%)	71 (88.8%)
nlink-out	84 (71.8%)	0 (0.0%)	83 (70.9%)	115 (98.3%)	83 (70.9%)	94 (80.3%)	84 (71.8%)	113 (96.6%)
nlink-all	1 (1.25%)	46 (57.5%)	0 (0.0%)	80 (100%)	17 (21.2%)	57 (71.2%)	23 (28.7%)	71 (88.8%)
nedit	105 (100%)	103 (98.1%)	105 (100%)	0 (0.0%)	105 (100%)	105 (100%)	98 (93.3%)	92 (87.6%)
pagerank	17 (21.2%)	46 (57.5%)	17 (21.2%)	80 (100%)	0 (0.0%)	48 (60.0%)	15 (18.8%)	68 (85.0%)
proposed-w	84 (77.8%)	85 (78.7%)	85 (78.7%)	108 (100%)	76 (70.4%)	0 (0.0%)	82 (75.9%)	99 (91.7%)
proposed-qv	43 (43.4%)	66 (66.7%)	42 (42.4%)	92 (92.9%)	34 (34.3%)	73 (73.7%)	0 (0.0%)	53 (53.5%)
proposed	71 (88.8%)	76 (95.0%)	71 (88.8%)	67 (83.8%)	68 (85.0%)	71 (88.8%)	34 (42.5%)	0 (0.0%)

目的として，編集履歴やリンク関係を用いる手法の提案及び，単純なアイデアを基にしたベースライン手法との比較実験を行った．実験より，更新エンティティの検出数の観点では比較手法の方が良く，非定常性スコアのもとでは提案手法が良い結果となることがわかった．また，各手法において実際に更新が起きると予測されたエンティティの実例を見てみると，リンク数を用いる手法が人の目につきやすい上位概念のエンティティを推薦するのに対し，提案手法は，直近の編集数などをもとに比較的マイナーなエンティティを推薦していることがわかった．これより提案手法は，目的であった実際のメンテナンスにおいて編集が見落とされるようなケースを防ぐのに利用できると考えられる．

次に，更新されたエンティティの周辺で，関連して更新が行われる可能性が高いエンティティを列挙する実験においては，リンク数を用いる手法や PageRank が，元のエンティティに追従して編集が行われる可能性の低い，上位概念のものを多く推薦する一方，提案手法では，人物，都市の両方のケースにおいて，関連して編集を行うべきと言えるようなものを多く推薦することができていた．このように，検出数の観点では単純な手法が良く，予測できた編集の非定常性や実際の推薦結果の観点では提案手法が良い結果となった原因としては，Wikidata においてエンティティに行われる編集の多くが，現実世界の事実の更新に基づくものでなく，管理のためのリファクタリング的な編集であることや，本来行われるべき編集がされていないことなどが考えられる．

また，今回提案した手法における隣接行列及びクエリベクトルそれぞれの有効性の観点では，直近の編集数を反映しているクエリベクトルに比較的強く影響を受けている一方，隣接行列の設定により検出数が向上したことから，両方を用いることでより情報が古いエンティティを減らすことができると考えられる．一方，周辺エンティティの推薦の結果をみると，提案手法はクエリベクトルのみを設定した場合と似通ったエンティティを高スコアとしていたことから，実用性の観点をもとに，調整を行う必要があると言える．

今後の展望としては，今回提案した手法を実際のシステムに導入することで，メンテナンスの絶対数やカバレッジを向上できるかなど，有用性についてのより正確な実験を行うとともに，その結果をもとにした手法改善を目指す予定である．

謝　　辞

本研究の結果は，SKY 株式会社 (C3I03115) による共同研究経費の助成及び国立研究開発法人新エネルギー・産業技術総合開発機構（NEDO）の委託業務（JPNP20006）の下で得られたものです.

文　　献

[1] C. Niklaus, M. Cetto, A. Freitas, and S. Handschuh, “A survey on open information extraction,” in *Proceedings of the 27th International Conference on Computational Linguistics*, (Santa Fe, New Mexico, USA), pp. 3866–3878, Association for Computational Linguistics, Aug. 2018.

[2] V. Yadav and S. Bethard, “A survey on recent advances in named entity recognition from deep learning models,” in *Proceedings of the 27th International Conference on Computational Linguistics*, (Santa Fe, New Mexico, USA), pp. 2145–2158, Association for Computational Linguistics, Aug. 2018.

[3] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, “Translating embeddings for modeling multi-relational data,” in *Advances in Neural Information Processing Systems* (C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, eds.), vol. 26, pp. 2787–2795, Curran Associates, Inc., 2013.

[4] B. Shi and T. Weninger, “Proje: Embedding projection for knowledge graph completion,” *CoRR*, vol. abs/1611.05425, 2016.

[5] R. Xie, Z. Liu, J. Jia, H. Luan, and M. Sun, “Representation learning of knowledge graphs with entity descriptions,” in *Proceedings of the Thirtieth AAAI Conference on Artificial*

表 6　各手法によって推薦されたエンティティ Q6279 (Joe Biden) の周辺の更新可能性が高いエンティティ

nlink-in		**nlink-out**	
relation	*tail*	*relation*	*tail*
work location	Washington, D.C.	different from	Joe Biden
different from	Joe Biden	depicted by	Joe Biden
residence	Claymont	notable work	Promises to Keep
member of political party	Democratic Party	spouse	Jill Biden
topic's main category	Category: Joe Biden	candidacy in election	2020 United States presidential election
nlink-all		**nedit**	
relation	*tail*	*relation*	*tail*
different from	Joe Biden	significant event	Joe Biden presidential campaign, 2020
depicted by	Joe Biden	sibling	Francis Biden
notable work	Promises to Keep	child	Naomi Biden
spouse	Jill Biden	child	Ashley Biden
candidacy in election	2020 United States presidential election	sibling	Valerie Biden Owens
pagerank		**proposed-w**	
relation	*tail*	*relation*	*tail*
significant event	Joe Biden presidential campaign, 1988	topic's main category	Category:Joe Biden
place of birth	St. Mary's Hospital	significant event	Joe Biden presidential campaign, 1988
residence	Arden	topic's main template	Template:Joe Biden
residence	Claymont	member of sports team	Delaware Fightin' Blue Hens football
notable work	Promises to Keep	notable work	Promises to Keep
proposed-qv		**proposed**	
relation	*tail*	*relation*	*tail*
child	Beau Biden	child	Beau Biden
significant event	Joe Biden presidential campaign, 2020	significant event	Joe Biden presidential campaign, 2020
sibling	Francis Biden	member of sports team	Delaware Fightin' Blue Hens football
award received	Grand Cross of the Order of Boyacá	child	Naomi Biden
child	Naomi Biden	sibling	Valerie Biden Owens

表 7　各手法によって推薦されたエンティティ Q1490 (Tokyo) の周辺の更新可能性が高いエンティティ

nlink-in		**nlink-out**	
relation	*tail*	*relation*	*tail*
named after	capital	head of government	Shunichi Suzuki
contains administrative territory	Chūō-ku	head of government	Ryokichi Minobe
〃	Musashino	flag	flag of Tokyo
〃	Shinjuku-ku	head of government	Ryotaro Azuma
shares border with	Chiba Prefecture	head of government	Seiichirō Yasui
nlink-all		**nedit**	
relation	*tail*	*relation*	*tail*
head of government	Shunichi Suzuki	contains administrative territory	Utsuki
head of government	Ryokichi Minobe	public holiday	Tokyo Citizen's Day
flag	flag of Tokyo	contains administrative territory	Ōkagō
head of government	Ryotaro Azuma	〃	Itsukaichi
head of government	Seiichirō Yasui	〃	Hōya
pagerank		**proposed-w**	
relation	*tail*	*relation*	*tail*
anthem	Tokyo Metropolitan Song	topic's main Wikimedia portal	Portal:Tokyo
public holiday	Tokyo Citizen's Day	public holiday	Tokyo Citizen's Day
category of people buried here	Catégorie: Personnalité inhumée à Tokyo	highest point	Mount Kumotori
flag	flag of Tokyo	economy of topic	economy of Tokyo
category for the view of the item	Category: Views of Tokyo	seal description	Symbols of Tokyo
proposed-qv		**proposed**	
relation	*tail*	*relation*	*tail*
public holiday	Tokyo Citizen's Day	public holiday	Tokyo Citizen's Day
headquarters location	Tokyo Metropolitan Government Complex	headquarters location	Tokyo Metropolitan Government Complex
archives at	Tokyo Metropolitan Archives	archives at	Tokyo Metropolitan Archives
contains administrative territory	Utsuki	contains administrative territory	Utsuki
history of topic	timeline of Tokyo	history of topic	timeline of Tokyo

Intelligence, AAAI'16, p. 2659–2665, AAAI Press, 2016.

[6] D. T. Wijaya, N. Nakashole, and T. M. Mitchell, "CTPs: Contextual temporal profiles for time scoping facts using state change detection," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (Doha, Qatar), pp. 1930–1936, Association for Computational Linguistics, Oct. 2014.

[7] J. Leblay and M. W. Chekol, "Deriving validity time in knowledge graph," in *Companion Proceedings of the The Web Conference 2018*, WWW '18, (Republic and Canton of Geneva, CHE), p. 1771–1776, International World Wide Web Conferences Steering Committee, 2018.

[8] G. Jeh and J. Widom, "Scaling personalized web search," in *Proceedings of the 12th International Conference on World Wide Web*, WWW '03, (New York, NY, USA), p. 271–279, Association for Computing Machinery, 2003.

[9] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," in *Proceedings of the 7th International World Wide Web Conference*, (Brisbane, Australia), pp. 161–172, 1998.

[10] D. Vrandečić and M. Krötzsch, "Wikidata: A free collaborative knowledgebase," *Communications of the ACM*, vol. 57, no. 10, pp. 78–85, 2014.

[11] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," *CoRR*, vol. abs/1907.10902, 2019.