

2025 年 7 月 24 日
データベースと応用システムレポート

7 月 18 日課題

情報経営システム工学分野 B4

学籍番号 : 24336488

氏名 : 本間三暉

問 20 情報検索の代表的なアルゴリズムである

HITS(Hyperlink-Induced Topic Search) と PageRank の共通点を説明しなさい。

HITS と PageRank はどちらも Web のハイパーリンク構造を解析し、重要度を数値化するリンク解析アルゴリズムである^[1]。共通点は以下の三点に整理できる。

- **リンクグラフ依存**：ページを頂点，リンクを有向辺と見なしてスコアを計算する^{[2] [3]}。
- **反復的固有ベクトル計算**：隣接行列に対してパワーイテレーションを行い，固有ベクトル（確率ベクトル）を収束させる^[3]。
- **再帰的スコア定義**：「重要なページから参照されるページは重要」という自己参照的な定義で評価値を決める^{[2] [3]}。

これらの特性により，両アルゴリズムは Web 規模でもスケーラブルにページ重要度を算出できる。

問 21 推薦システムの代表的な手法である協調フィルタリングの利点と欠点を説明しなさい。

利点

- **コンテンツ独立性**：アイテム属性を解析せず，ユーザ - アイテム行列のみで機能する^[4]。
- **高精度なパーソナライズ**：潜在的な嗜好パターンを学習し，複雑な好みも捉えやすい^[4]。
- **ドメイン横断適用**：映画，音楽，EC 商品など多様な領域で実績がある^[5]。
- **セレンディピティ**：ユーザ自身が気付かなかった関連アイテムを提示しやすい^[5]。

欠点

- **コールドスタート問題**：履歴のない新規ユーザ・アイテムで精度が下がる^[4]。
- **スパース行列問題**：大規模カタログでは評価行列が疎で類似計算が難しい^[4]。
- **計算コスト増大**：ユーザ数やアイテム数が増えると計算量が増大する^[5]。
- **人気バイアス／フィルターバブル**：人気アイテムに推薦が偏り，多様性が失われる^[5]。
- **シリング攻撃の脆弱性**：悪意ある評価により推薦結果がゆがむ恐れがある^[4]。

以上の利点と欠点を踏まえ，実システムではハイブリッド手法や正則化などを組み合わせて精度と健全性を両立させることが求められる。

参考文献

[1] データベースと応用システムの講義資料

[2] 分析ノート「HITS アルゴリズム」

<https://analytics-note.xyz/graph-theory/hits-algorithm/>

[3] け日記「いまさら学ぶ PageRank アルゴリズム」

<https://ohke.hateblo.jp/entry/2018/12/29/230000>

- [4] Udeemy メディア「協調フィルタリングって何？商品のおすすめ機能を学ぼう！」

<https://udemy.benesse.co.jp/data-science/ai/collaborative-filtering.html>

- [5] FeliCa Networks「協調フィルタリングとは？基本的な考え方や種類を解説」

<https://receiptreward.jp/solution/column/collaborativefiltering.html>