

Web ページ推薦システムにおけるユーザ指向型スニペット生成

渡邊 亮彦[†] 笹野 遼平[‡] 高村 大也[‡] 奥村 学[‡]

[†] 東京工業大学 総合理工学研究科, [‡] 東京工業大学 精密工学研究所

[†]watanabe@lr.pi.titech.ac.jp, [‡]{sasano, takamura, oku}@pi.titech.ac.jp

1 はじめに

近年、ブログやSNSに代表されるCGMの普及に伴い、Web上の情報が爆発的に増大し、個人にとって有用な情報を取得するためには膨大な労力が必要となってきた。この問題を解決するために、様々なWebページ推薦システムが研究されてきた。

Webページ推薦システムは、個々のユーザの嗜好に基づいて自動的に好みのページを推薦する。ユーザがシステムを利用する際は、推薦結果の一覧から興味のあるページを探す必要がある。これを支援するために、システムは推薦結果の一覧に各推薦ページのタイトルやスニペット（ページの概要文）などの情報を付与するのが一般的である。しかし推薦システムでは検索キーワードが存在しないため、検索エンジンと同様の手法でスニペットを生成することはできない。また、ユーザごとに欲する情報は異なるため、提示するスニペットを個々のユーザの嗜好に適応させる必要がある。

そこで本研究では、Webページ推薦システムによって推薦された各ページに対してユーザ指向型のスニペットを生成する手法を提案する。具体的には、Webページ推薦システムが推薦を生成した際の推薦理由を直接的に利用することでスニペットを生成する。これにより、スニペットに個々のユーザの嗜好を反映するだけでなく、各ページが推薦されるに至った理由を含むような記述を含めることが可能であると考えられる。

2 関連研究

スニペット生成手法の一つとして、文書要約を用いる手法が提案されている。その中でも、Personalized Summarization（以下、PS）が関連深い。PSはユーザの嗜好情報を利用して個々のユーザに特化した文書要約を行う研究課題である。

Quら[1]は、ソーシャルタギングシステムにおけるタグ推薦と単一文書要約の手法を相互作用させることによりPSを実現している。本研究とは、推薦シ

ステムの内部情報を用いてPSを行う点で類似している。しかしタグ推薦の情報を利用するだけでは、要約文中にページが推薦される動機となった情報を十分に含めることは困難である。本研究ではWebページ推薦を行った際の推薦理由を直接的に単語重要度の決定に利用しスニペット生成に寄与させることで、推薦の動機となった情報をスニペット中により反映させることを試みる。また、Quらの手法はベースとしているタグ推薦の手法に依存しているため汎用性が低いのに対して、本研究は推薦理由を用いて単語重要度を決定することができればさまざまな推薦手法に対して適用可能であることは特筆すべき点である。

Díazら[2]は、コンテンツベースのニュース記事推薦システムの内部情報を用いることでPSを行なっている。このシステムではユーザの長期的嗜好と短期的嗜好を収集し、それらの情報に基づいてニュース記事の推薦を行う。長期的嗜好は、予めユーザがニュースのカテゴリやキーワードなどの入力を行うことで収集され、短期的嗜好は適合性フィードバックを用いて収集される。PSを行う際には、これらの推薦システムが収集した嗜好情報を用いることで、個々のユーザの興味に沿った要約を生成する。本研究とは推薦システムの内部情報を利用して、PSを行う点で類似している。しかし、本研究ではユーザに明示的な嗜好情報の入力を求めることはせず、暗黙的に獲得された嗜好情報のみを用いてスニペットを生成する点で異なる。

3 提案手法

提案手法ではまずユーザに対して推薦するWebページの決定を行い、その後個々の推薦ページに対してユーザ指向型スニペットを生成する。このとき、推薦理由を用いて単語重要度を決定し文書要約に活用することでユーザ指向型スニペット生成を行う。提案手法は推薦理由を用いて単語重要度を決定することができれば、さまざまな推薦手法において汎用的に利用可能である。

3.1 Web ページの推薦

本研究の主眼はユーザ指向型スニペット生成であるため、Web ページ推薦手法は既存の手法、具体的にはソーシャルブックマーク（以下 SBM）を利用する Zhang ら [3] の手法を用いる。SBM は Web 上で他者と自身のブックマークを共有できるサービスであり、各ブックマークにはタグと呼ばれるキーワードを付与できる。SBM を利用した Web ページ推薦では、これらのユーザが登録したブックマークやタグに関する情報を嗜好情報として利用し、Web ページ推薦を行う。

Zhang らの手法では、ユーザスコアとタグスコアに基づいて個々の推薦ページが決定される。以下、ユーザスコアとタグスコアを推薦理由の観点から説明する。

ユーザスコア $u_score(u_t, u_l, p_n)$

被推薦ユーザ u_t に対して推薦を生成したときに、あるユーザ u_l がどれだけ推薦ページ p_n の推薦に寄与したかを表す。

タグスコア $t_score(u_t, t_m, p_n)$

被推薦ユーザ u_t に対して推薦を生成したときに、あるタグ t_m がどれだけ推薦ページ p_n の推薦に寄与したかを表す。

上記のユーザスコアとタグスコアは、スニペットを生成する際に用いる単語重要度の決定に活用する。

3.2 ユーザ指向型スニペット生成

3.2.1 文書要約モデル

スニペットの生成には、文選択による文書要約モデルを用いる。具体的には文書要約を最大被覆問題として定式化し、重要度の高い単語をより多く被覆した要約を生成することで高い性能を示した高村ら [4] の手法を用いる。提案手法では高村らのモデルに被推薦ユーザ u_t の単語 e_j に対する単語重要度 w_{tj} を導入することで、ユーザ指向型スニペット生成を実現する。したがって、本研究が扱う文書要約モデルは、以下のような整数計画問題として表現される：

$$\begin{aligned} & \text{maximize } (1 - \lambda) \sum_j w_{tj} z_j + \lambda \sum_i \left(\sum_j w_{tj} a_{ij} \right) x_i \\ & \text{s.t. } \sum_i c_i x_i \leq K; \quad \forall j, \sum_i a_{ij} x_i \geq z_j; \\ & \quad \forall i, x_i \in \{0, 1\}; \quad \forall j, z_j \in \{0, 1\}. \end{aligned}$$

ここで、 x_i は文 s_i が選択された場合に 1、そうでない場合は 0 となる変数であり、 z_j は単語 e_j が被覆さ

れた場合に 1、そうでない場合は 0 となる変数である。また、 a_{ij} は文 s_i 内に単語 e_j が含まれている場合に 1 となり、そうでない場合は 0 となる定数であり、 c_i は文 s_i の文字数、 K はスニペット長を表す定数である。 λ は生成するスニペットの冗長性を制御する定数であり、 λ の値が小さいほど冗長性を排除したスニペットが生成される。制約条件の $\sum_i c_i x_i \leq K$ は生成したスニペットの長さがスニペット長 K 以下に収まるようにする制約である。また、制約条件の $\sum_i a_{ij} x_i \geq z_j$ は単語 e_j の被覆に関する制約を表している。すなわち、単語 e_j が含まれている文が一つ以上選択されたときに、 e_j は被覆されたとみなす。これにより、推薦理由を反映する単語をより多く被覆したスニペットが生成可能となる。

高村らの手法はもともと複数文書要約のために提案された手法であることに注意する。本研究が扱うタスクは単一文書要約とみなすことができるため、上記のモデルを単一文書要約に適した形に拡張する。具体的には、ある文書が与えられたとき、文書の先頭に近い文ほど重要な情報が含まれていることが多いことに着目し、以下のようにモデルの目的関数を拡張する：

$$\text{maximize } (1 - \lambda) \sum_j w_{tj} z_j + \lambda \sum_i \frac{\left(\sum_j w_{tj} a_{ij} \right)}{\text{pos}(s_i)} x_i.$$

ここで $\text{pos}(s_i)$ は文 s_i が文書中に出現する位置を表す。また制約条件は同様である。これにより先頭に近い文ほど重要度が高くなりスニペットに含まれやすくなる。

3.2.2 単語重要度 w_{tj} の決定

単語重要度 w_{tj} の決定における基本的な考え方は、「推薦ページに対するスコアが高いユーザやタグの特徴語はより重要である」というものである。このような考え方に基づいて単語重要度を決定するために、本研究では *TF-IDF* を活用する。具体的には、ある推薦ページ p_n が与えられたとき、以下のような式に基づき w_{tj} を決定する：

$$\begin{aligned} w_{tj} = & \sum_l u_score(u_t, u_l, p_n) \times TF(u_l, e_j) \times IDF(e_j) \\ & + \sum_m t_score(u_t, t_m, p_n) \times TF(t_m, e_j) \times IDF(e_j). \end{aligned}$$

ここで、

$$TF(u_l, e_j) = \frac{u_l \text{の全ブックマークページ中の単語 } e_j \text{ の出現回数}}{u_l \text{の全ブックマークページ中の総単語数}}$$

$$TF(t_m, e_j) = \frac{t_m \text{が付与されている全ページ中の単語 } e_j \text{ の出現回数}}{t_m \text{が付与されている全ページ中の総単語数}}$$

$$IDF(e_j) = \log \left(\frac{\text{総ページ数}}{\text{単語 } e_j \text{ が出現したページ数}} \right)$$

である。これにより、推薦ページに対してスコアが高いユーザやタグの特徴語により高い重要度を割り当てることで、単語重要度に推薦に至った動機に関する情報を反映させることができる。このように、ベースとする推薦手法の推薦理由を用いて単語重要度 w_{tj} の決定を行うことができれば、さまざまな推薦手法に対して提案手法は適用できることを改めて強調しておく。

4 実験

4.1 データセット

本実験では推薦用データセットとスニペット評価用データセットの2種類のデータセットを用いる。推薦用データセットは、ユーザに対して推薦を生成する際に用いるデータセットである。またスニペット評価用データセットは、ユーザに対してWebページの推薦を行い、推薦されたページに対してユーザ自身に正解となるスニペットを生成してもらうことで作成したデータセットである。

4.1.1 推薦用データセット

本実験では推薦用データセットとして livedoor クリップデータセット¹を用いた。提案手法はニュース記事のような簡潔な記事ではなく、ブログ記事のような、ある程度長文で雑多な内容を含む記事においてより有効に働くと考えられる。そこで、今回は livedoor クリップデータセットのうち、主要な17種類のブログサイトに属するエントリを抽出し、さらにスパースなデータセットにならないよう以下の処理を施し推薦用データセットとした。

- ブックマーク数が5件以上、かつ利用したタグの異なり数が5つ以上のユーザを抽出
- 被ブックマーク数が5件以上、かつ付与されたタグの異なり数が5つ以上のWebページを抽出
- 利用したユーザ数が5人以上、かつ付与されたWebページ数が5件以上のタグを抽出

最終的に推薦用データセットは、2,458人のユーザ、10,501個のWebページ、4,027個のユニークタグによって構成された。また、ユーザのブックマークの合計は101,792件、タグ付け数の合計は227,025件であった。

¹現在は公開を終了している

4.1.2 スニペット評価用データセット

ユーザごとに欲しいスニペットは異なると考えられるため、提案手法の評価には、ユーザごとに個別に作成したスニペットが必要となる。しかし提案手法を評価するための代表的なデータセットがないため、11人のユーザに協力してもらい、以下の(1)、(2)、(3)に従いスニペット評価用データセットの作成を行った。

(1) 嗜好情報の収集

ユーザに対してWebページを推薦するためには、ユーザの嗜好情報が必要となる。そのため、今回はユーザに対してWebページを提示し、提示されたページのうち、ユーザが気に入ったページに対してブックマーク登録とタグ付けを行ってもらった。すなわち、ユーザのページに対するブックマークとタグ付けに関する情報が嗜好情報となる。このとき、ユーザには最低15件のブックマーク登録を行ってもらい、ブックマークしたそれぞれの記事に対してタグ付けを行ってもらった。提示するページは推薦用データセットに含まれるものとした。

(2) Webページの推薦

Zhangら[3]の手法を用いて(1)で収集した嗜好情報に基づきユーザに対してWebページの推薦を行った。このときユーザごとに10件のページを推薦した。

(3) ユーザによるスニペットの生成

(2)でユーザに推薦したWebページのうち、ユーザが興味のあるページに対して、自身が欲しいと考えるスニペットを生成してもらった。ユーザにスニペットを生成してもらう際には、Webページの本文のうち、自身がスニペットとして欲しい文を150文字未満で選択してもらった。このとき、「スニペット全体を読むことで、自身が該当ページに興味があることを適切に判断できる」ように文を選択するように指示した。最終的に、11人のユーザに推薦されたページのうち、合計50個のページに対する正解スニペットを得たため、これらをスニペット評価用データセットとした。

4.2 評価尺度

評価尺度としては、ROUGE-2[5]を用いた。ROUGE-2は単一文書要約の評価において、人間の評価と非常に強い相関を示すことが報告されている[5]。本研究のタスクも単一文書要約とみなせることより、ROUGE-2を評価尺度として採用した。ROUGE-2を用いることで、生成したスニペットが正解スニペットに含まれる単語バイグラムをどれだけ被覆できたかを評価することが可能である。

4.3 実験設定

実験のベースラインとして、リード文、すなわち Web ページ中の本文を先頭から逐次抽出しスニペットとする手法（リード文）、Collaborative Summarization (CS) [1]、提案手法において単語重要度の決定に推薦理由を用いない手法（Generic）の3種類を用いた。スニペットの長さは150文字とした。また、Genericで単語重要度の決定を行う際には *TF-IDF* を使用した。このとき、*TF* はスニペットを生成するページ内に含まれる単語の頻度から算出した。

スニペット評価用データセットに対して leave-one-out 交差検定を行った。すなわち、スニペット評価用データセット50件のうち1件をテスト事例、残りの49件を訓練事例とし、これらを50件すべての事例が1回ずつテスト事例となるように評価を行った。各手法のパラメータの調整は、訓練事例においてストップワードを除去しない場合の ROUGE-2 スコアの平均が最大となるように行った。ROUGE-2 は単語バイグラムの被覆率を表すスコアであるため、ストップワードの除去を行ってしまった場合、ROUGE-2 スコアに寄与する単語バイグラムが正解スニペットの本来の単語バイグラムに準拠しないことに留意した。

4.4 実験結果および考察

表1に ROUGE-2 スコアによる各手法の比較を示す。ここで、表1の提案手法（拡張）は文の位置情報を考慮したモデルを表し、提案手法は文の位置情報を考慮しないモデルを表す。また、表1に示したスコアはストップワードの除去を行っていない場合の ROUGE-2 スコアの平均である。

提案手法、Generic、CS はそれぞれ文の位置情報を考慮しないモデルであるが、文の位置情報を考慮する提案手法（拡張）およびリード文と比較して ROUGE-2 スコアが大幅に低いことがわかる。これは、Web ページの先頭に近い文ほどスニペットに含めるべき文が多い傾向があることを示している。一方、提案手法（拡張）とリード文の結果を比較すると、提案手法（拡張）の ROUGE-2 スコアはリード文を0.032ほど上回っていることがわかる。これはページの先頭にある文だけでなく、推薦理由を反映した文を冗長性とのバランスを考慮しながらスニペットに含めることで ROUGE-2 スコアの上昇に寄与していることを示している。また、提案手法と Generic、CS を比較すると提案手法の ROUGE-2 スコアはそれぞれ0.019、0.029ほど上回っていることがわかる。これは、推薦理由を用いた

表 1: ROUGE-2 スコアによる各手法の比較

手法	ROUGE-2
提案手法（拡張）	0.277
リード文	0.245
提案手法	0.132
Generic	0.113
CS	0.103

単語重要度の決定が、ROUGE-2 スコアの上昇に寄与していることを示している。

5 おわりに

Web ページ推薦システムにおいて推薦されたページに対して、ユーザ指向型スニペットを生成する手法を提案した。提案手法では推薦システムの推薦理由を用いて単語重要度を決定し文書要約モデルに導入することでユーザ指向型スニペットを生成した。

今回の評価では正解スニペットを生成したページはユーザが興味のあるページだけに留まってしまっていることに注意する必要がある。現状の評価では提案手法によって生成されたスニペットを用いることで、ユーザが興味がないページを適切に興味がないと判断できるか否かという観点が欠落している。そのため、検索エンジンにおけるスニペットをタスクに基づいて評価するのと [6] 同様に、より実際の運用に近い形で評価することが今後の課題として挙げられる。

参考文献

- [1] Yang Qu and Qunxiu Chen. Collaborative Summarization: When Collaborative Filtering Meets Document Summarization. *PACLIC 2009*, pp. 474–483, (2009).
- [2] Alberto Díaz and Pablo Gerás. User-model based personalized summarization. *Information Processing and Management*, 43(6), pp. 1715–1734 (2007).
- [3] Zi-Ke Zhang, Tao Zhou, and Yi-Cheng Zhang. Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs. *Physica A* 2010, 389(1), pp. 179–186, (2010).
- [4] 高村大也, 奥村学. 最大被覆問題とその変種による文書要約モデル. 人口知能学会論文誌, 23(6), pp. 505–513, (2008).
- [5] Chin-Yew Lin. ROUGE: A Package for Automatic Evaluation of Summaries. *ACL 2004*, pp. 74–81, (2004).
- [6] Anastasios Tombros and Mark Sanderson. Advantages of Query Biased Summaries in Information Retrieval. *ACM-SIGIR 1998*, pp. 2–10, (1998).