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

Generating Personalized Snippets for Web Page Recommender Systems

渡邊 亮彦
Akihiro Watanabe

東京工業大学
Tokyo Institute of Technology
watanabe@lr.pi.titech.ac.jp, <http://www.lr.pi.titech.ac.jp/~watanabe/>

笹野 遼平
Ryohei Sasano

東京工業大学 科学技術創成研究院 未来産業技術研究所
Tokyo Institute of Technology, Institute of Innovative Research, Laboratory for Future Interdisciplinary Research of Science and Technology
sasano@pi.titech.ac.jp, <http://www.lr.pi.titech.ac.jp/~sasano/>

高村 大也
Hiroya Takamura

(同 上)
takamura@pi.titech.ac.jp, <http://www.lr.pi.titech.ac.jp/~takamura/>

奥村 学
Manabu Okumura

(同 上)
oku@pi.titech.ac.jp, <http://www.lr.pi.titech.ac.jp/~oku/>

keywords: personalized snippets, recommender systems, relevance judgment

Summary

Web page recommender systems usually provide users with titles and snippets of recommended pages when the systems present a list of recommendations. Snippets help users judge whether recommended web pages are relevant or not. However, while search engines usually show a text span around a search query as a snippet, web page recommender systems cannot leverage the snippet generation methods used by search engines because the recommender systems have no search queries. Web page recommender systems thus generally use lead sentences, i.e. the first sentences of web pages, as a snippet, but lead sentences are not necessarily relevant to user's interest. Furthermore since user's information needs can be different from each other, personalized snippets are desirable to support user's relevance judgment. Therefore, we propose a new method to generate personalized snippets for web page recommender systems that uses reasons why the web pages are recommended to the user. This use of reasons enables snippets to reflect the interest of the user. Furthermore, since our formulation does not depend on a certain recommender system, our method can be applied to diverse recommender systems. The experimental result on manually created dataset shows that our method is superior to the existing method and generic summarization model in terms of ROUGE-2. In addition, our method achieves comparable performance with the lead method despite that our method restricts itself to sentence selection while the lead method is free to extract a part of a sentence at the end of its snippets.

1. は じ め に

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システムであり、近年注目が高まっている研究分野である [Park 12]。Web ページ推薦システムがユーザに推薦結果を提示する際は、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）を付与するのが一般的である。これらの情報は、ユーザが推薦されたページを閲覧するか否かを判断する（以下、適合性判定）ために提供される。特に、スニペットは検索エンジンで採用されており、Turpin ら [Turpin 09] をはじめとする多くの研究で、適合性判定における重要性が明らかにされている。そのため、スニペットは Web ページ推薦システムにおいても有用であると考えられる。本稿では、Web ページ推薦システムにおいて

スニペットを生成する手法を取り扱う。

スニペットには、適合性判定に寄与する情報がより多く含まれることが望ましい。検索エンジンにおいては、ユーザから明示的に検索クエリが入力されるため、検索クエリの近傍にあるテキストをスニペットとして提示することで、適合性判定を支援している。一方、Web ページ推薦システムでは、ユーザからの明示的な情報の入力を前提としないものが一般的である^{*1}。このようなシステムの場合、ソーシャルメディアなどにおけるユーザの活動から暗黙的な嗜好情報を獲得し、推薦を行うシステムが主流である。これは、推薦を行うためには多くの情報が必要である一方で、それらを明示的にユーザに入力してもらうのは負担が大きいだけでなく、そもそもどのよ

^{*1} コンテンツベースの推薦システムなどでは、推薦のためのキーワードを明示的に入力するものも存在する。

うな入力をすれば満足のいく推薦を得られるか分からないからである^{*2}。そのため、検索エンジンと同様の仕組みでスニペットを生成するのは困難であり、多くの Web ページ推薦システムでは、リード文（ページの冒頭文）をスニペットとしている。しかしながら、ユーザごとに求める情報は異なるため、ユーザが適切に適合性判定を行うためには、個々のユーザの嗜好を反映したスニペットが望ましい。そこで、我々はユーザの嗜好を手掛かりにスニペットを生成することを考える。たとえば、図 1 左に示したようなプロ野球の速報ページが、阪神ファンと巨人ファンのユーザに推薦されたと仮定する。阪神ファンと巨人ファンであれば、それぞれのチームの試合結果の速報には興味があり、それらの情報は適合性判定に寄与すると考えられる。したがって、図 1 右に示すように、巨人ファンには巨人の試合結果の情報、阪神ファンには阪神の試合結果の情報を含むようなスニペットを提示することで適合性判定を支援できると考える。本研究では、このようにユーザの嗜好に合わせて内容が動的に変化するスニペット、すなわち、ユーザ指向型スニペットを生成するモデルを提案する。

文書要約において、ユーザの嗜好に合わせて個々のユーザに特化した要約を作成する研究は、いくつか行われている。これらの研究は、ユーザとシステムのインタラクションを通じて、明示的に獲得されたユーザの嗜好情報を用いる研究 [Díaz 07, Yan 11, Zhang 03]、システムが自動的に獲得したユーザの嗜好情報を用いる研究 [Qu 09] に大別される。

本研究では、基本的に自動的に獲得された嗜好情報を用いる推薦システムでの利用を想定しているため、Qu ら [Qu 09] と同様に、後者のアプローチに焦点を当てる。Qu らの手法では、FolkRank [Hotho 06] の利用を想定している。一方、本研究では、特定の推薦システムを想定しておらず、多くのシステムに適用可能な汎用的な枠組みを提案する。また、提案手法では、同一ユーザの様々な嗜好情報に基づくスコアを用いて、統合的に文の重要度を算出し要約を作成することができる。たとえば、Qu らの手法では、FolkRank によって求まるタグのスコアを用いて要約を作成しているのに対し、本研究が提案する枠組みでは、タグのみならず、近傍ユーザや近傍ページなどの情報も用いて要約を作成することができる。このような、近傍ユーザや、近傍ページなど、推薦システムにおいてシステムが内部で保持している推薦の根拠となった情報のことを、本研究では「推薦根拠」と呼ぶ。

提案手法の基本的な戦略は、ユーザが興味を持つ情報と関連する単語は重要であると考え、それらの単語をより多く被覆するようなスニペットを生成することである。そのために、提案手法では Web ページ推薦システムの推薦根拠を活用するモデルを取り扱う。具体的には、推

阪神 試合結果	今日の交流戦では、計6試合が行われた。 「阪神 9-5 ソフトバンク」 阪神が連敗を3で止め、今季交流戦初勝利。1回西岡の自身18本目の先頭打者弾となる1号ソロで先制。更に新井良が自身初の満塁弾となる4号を放ち、自身最多の4打点。今季2度目の初回5得点。6回鳥谷が4号2ランで、今季最多タイ ...	😊 阪神ファンに対するスニペット 「阪神 9-5 ソフトバンク」 阪神が連敗を3で止め、今季交流戦初勝利。1回西岡の自身18本目の先頭打者弾となる1号ソロで先制。更に新井良が自身初の満塁弾と ...
巨人 試合結果	「巨人 2-1 西武」 巨人が延長10回相手の失策で今季4度目のサヨナラ勝ち。今季5度目3連勝で最多タイ貯金14。阿部が今季2度目、自身3度目の3試合連続12号ソロ。なお阿部は歴代単独3位の交流戦通算43本 ...	😊 巨人ファンに対するスニペット 「巨人 2-1 西武」 巨人が延長10回相手の失策で今季4度目のサヨナラ勝ち。今季5度目3連勝で最多タイ貯金14。阿部が今季2度目、自身3度目の3試合連続 ...

図 1 ユーザ指向型スニペットの例

薦根拠はユーザの嗜好を反映していると考えられるため、推薦根拠と関連が深い単語は重要であるとみなす。このようにして重要な単語を求めた後、文書要約モデルを適用することでスニペットの生成を行う。

推薦根拠を用いたモデル化を行う利点は三つある。一つ目は、推薦根拠を用いることで、推薦システムが保持する種々のユーザの嗜好に関する情報を、統合的に扱うことが可能となる点である。これにより、同一ユーザの様々な嗜好情報から、単語の重みを算出することができる。二つ目は、推薦に寄与した個々のデータと、個々のデータが与えたスコアが推薦根拠として与えられれば、提案手法は様々な Web ページ推薦システムにおいて利用可能となる点である。推薦システムの基本的なアプローチは、個々のデータが推薦に寄与する度合いをスコアとして算出し、算出したスコアを集約することで推薦を生成するものである。そのため、多くの Web ページ推薦システムにおいて、前述した形式で推薦根拠を与えることが可能であると考えられる。三つ目は、推薦根拠が良質なものとなれば、生成されるユーザ指向型スニペットも良質なものとなることが期待できる点である。推薦アルゴリズムの進歩に伴い、推薦根拠は今後さらに良質になることが予想されるため、提案手法の性能の向上が見込まれる。

2. 関 連 研 究

Web ページ推薦システムの多くは CTS (Collaborative Tagging Systems)^{*3}を活用したものである。特に、グラフに基づく手法、テンソル分解に基づく手法、トピックモデルに基づく手法などが注目を集めている [Wetzker 09, Xu 06, Zhang 10]。これらの研究は主に推薦の質を改善することに注力しており、Web ページ推薦システムにおけるスニペットについては取り扱っていない。

情報検索の分野においては、query-biased な要約手法 [Tombros 98] が、検索エンジンのスニペットを生成する

^{*2} たとえば、推薦を行うためのキーワードを求められた場合、ユーザが必ずしも自分の嗜好を言語化できるとは限らない。

^{*3} 代表的なものとして、はてなブックマーク (<http://b.hatena.ne.jp/>)、Flickr (<https://www.flickr.com/>) などが挙げられる。

際に用いられている．query-biased な要約手法では，スニペットを生成する対象であるページの内容，すなわち単語の出現頻度や文のページ内での出現位置，クエリとの関連度，テキストの書式などの情報を用いる手法や [Tombros 98, White 03]，ユーザに提示した情報の読み時間を用いた暗黙的な適合性フィードバックを利用する手法 [White 02] がある．さらに，Wang ら [Wang 07] は，これらの情報を適切な重みで組み合わせるために，ランキング学習を用いる手法を提案した．このような検索エンジンにおける query-biased な要約手法では，ユーザからクエリが与えられることが前提となっている．一方，Web ページ推薦システムではクエリが存在しないため，これらの手法を適用することはできない．

本研究と関連が深い別の文書要約タスクとして，PDS (Personalized Document Summarization) がある．PDS はユーザの嗜好情報を利用し個々のユーザに特化した文書要約を行うタスクである．

Qu ら [Qu 09] は，CTS におけるタグ推薦の手法の一つである FolkRank [Hotho 06] と，グラフに基づく単一文書要約手法 [Wan 07] を相互作用させることにより PDS を実現している．Qu らの手法では，FolkRank によって求まるタグの重みを，文の重要度に反映させることで，文書要約を行っている．このことから，Qu らの手法は推薦システムが保持するタグの情報を用いて PDS を行う手法と解釈できる．一方，提案手法の枠組みでは，タグ情報の利用のみに限定されず，近傍ユーザや近傍ページなど，様々な嗜好情報を考慮して PDS を行うことができる．Qu らの手法は，ベースとしているタグ推薦手法に強く依存しているため，汎用性が低いものに対して，提案手法は，推薦に寄与した個々のデータと，個々のデータが推薦に寄与した度合いを推薦根拠として与えるだけで PDS を行えるため，特定の推薦システムに依存しない汎用的な手法である．

Díaz ら [Díaz 07] は，コンテンツベースのニュース記事推薦システムが持つユーザの嗜好情報を利用し PDS を行っている．Díaz らのシステムは，嗜好情報としてユーザが興味を持つニュースのキーワードなどを用いており，これらはユーザが明示的にシステムに入力する必要がある．他にも，ユーザとシステムのインタラクション^{*4}を通じて，明示的に嗜好情報を収集した上で PDS を行う研究も存在する [Yan 11, Zhang 03]．これらのシステムでは，ユーザから明示的に嗜好情報が入力されることを前提としている．一方，提案手法は，ユーザからの明示的な嗜好情報の入力を必要とせず，推薦システム内部で暗黙的に保持されている推薦根拠を用いて PDS を行う．

3. 代表的な推薦システムのアルゴリズム

提案手法の説明に先立ち，代表的な推薦システムのアルゴリズムについて述べる．以下，本稿ではアイテム^{*5}を推薦する対象のユーザをターゲットユーザと呼ぶ．

3.1 協調フィルタリングに基づく手法

協調フィルタリングは，ターゲットユーザがあるアイテムを好む度合いを予測する手法である．具体的には，ターゲットユーザ c にとって未知のアイテム i に対する評価値 $\hat{r}_{ci}$ を予測する問題として定式化される．このとき，ターゲットユーザとその他のユーザが過去にアイテムに与えた評価値の情報に基づいて，評価値の予測を行う．協調フィルタリングでは，各ユーザの嗜好情報として，ユーザおよびアイテム，ユーザがアイテムに与えた評価値の三つ組を考える．そして，ユーザを行，アイテムを列，評価値を要素として持つ評価値行列を作成し，評価値の予測に活用する．以下，協調フィルタリングの中でも代表的な，ユーザ間型メモリベース法 [Adomavicius 05] とアイテム間型メモリベース法 [Sarwar 01] について説明する．

§1 ユーザ間型メモリベース法

ユーザ間型メモリベース法 [Adomavicius 05] は，ターゲットユーザと似た嗜好を持つユーザが好むアイテムを推薦する手法である．基本的なアプローチとしては，ターゲットユーザ c と最も類似する N 人のユーザを求め，類似ユーザのアイテム i に対する評価値を集約することによって $\hat{r}_{ci}$ を予測する．集約の仕方には様々な方法があるが，一般的には，ユーザ間の類似度によって重み付けした，類似ユーザ c' のアイテム i に対する評価値 $r_{c'i}$ の加重平均を用いる手法が用いられる．具体的には，ターゲットユーザ c の，あるアイテム i に対する評価値 $\hat{r}_{ci}$ を，以下のように決定する：

$$\hat{r}_{ci} = \frac{1}{Z} \sum_{c' \in \hat{C}_{ci}} \text{sim}(c, c') \times r_{c'i}.$$

ここで， $\hat{C}_{ci}$ は，アイテム i を過去に評価したことがあるユーザのうち，ターゲットユーザ c と最も類似した N 人のユーザの集合である．また， Z は正規化定数であり， $Z = \sum_{c' \in \hat{C}_{ci}} |\text{sim}(c, c')|$ である．ユーザ間の類似度 $\text{sim}(c, c')$ は，ユーザ c および c' が，共通して評価しているアイテムの評価値に基づいて算出される．これは，同じアイテムに対して，似たような評価値を与えるユーザ同士の嗜好は類似しているという考えに基づいている．評価値行列の行ベクトルは，ユーザの各アイテムに対する評価値を表すベクトルであるため，行ベクトル間のコサイン類似度，あるいはピアソンの相関係数を計算することで，ユーザ間の類似度を算出する．

^{*5} 本稿では推薦システムがユーザに推薦するものを総称してアイテムと呼ぶ．商品推薦であれば商品，音楽推薦であれば曲，Web ページ推薦であれば Web ページがアイテムに相当する．

^{*4} テキストのハイライト，あるいは興味のある文の選択など．

§2 アイテム間型メモリベース法

アイテム間型メモリベース法 [Sarwar 01] は、ターゲットユーザが過去に好んだアイテムと類似したアイテムを推薦する手法である。ユーザ間型メモリベース法では、 $\hat{r}_{ci}$ を予測する際にユーザ間の類似度を用いたが、アイテム間型メモリベース法では、アイテム間の類似度を用いる。具体的には、 $\hat{r}_{ci}$ を求める際に、ターゲットユーザ c が過去に評価したアイテムのうち、アイテム i と最も類似するアイテムを N 個求め、アイテム間の類似度によって重み付けした、ターゲットユーザ c の類似アイテム i' に対する評価値 $r_{ci'}$ の加重平均を用いて $\hat{r}_{ci}$ を予測する。定式化すると、以下ようになる:

$$\hat{r}_{ci} = \frac{1}{Z} \sum_{i' \in \hat{I}_{ci}} \text{sim}(i, i') \times r_{ci'}.$$

ここで、 $\hat{I}_{ci}$ は、ターゲットユーザ c が過去に評価したアイテムのうち、アイテム i と最も類似した N 個のアイテムの集合である。また、 Z は正規化定数であり、 $Z = \sum_{i' \in \hat{I}_{ci}} |\text{sim}(i, i')|$ である。アイテム間の類似度 $\text{sim}(i, i')$ は、多くのユーザから似たような評価を受けるアイテム同士は類似しているという考えに基づいて算出される。すなわち、評価値行列の列ベクトル間の類似度を測ることにより算出される。

3.2 CTS に基づく手法

CTS に基づく推薦システムの代表的手法として、Zhang らの手法 [Zhang 10] を説明する。特に、Zhang らの手法は実験において用いるため、詳しく説明する。

CTS ではユーザ間でアイテムを共有し、自身の興味に基づいて新たなアイテムを見つけることができる。加えて、CTS ではユーザが新たなアイテムを発見し共有するのを手助けするために、タグと呼ばれるキーワードをアイテムに付与することができる [Golder 06]。CTS に基づく推薦システムでは、ユーザのこのような CTS 上での活動によって生じるユーザ、アイテム、あるいはタグ間の関係を活用して推薦を行う。特に Zhang らの手法は、CTS におけるこれらの関係をグラフで表現し、推薦スコアをグラフ上で伝搬させることで推薦を行う。ここで推薦スコアとは、あるユーザに対してアイテムを推薦するときの推薦度を表すスコアである。

§1 CTS のグラフ表現

CTS におけるユーザ、アイテム、タグ間の関係は、グラフを用いて表現することができる。特に Zhang らは、ユーザ-アイテム二部グラフと、アイテム-タグ二部グラフによって構成される三部グラフを用いて、CTS を表現した。

ユーザの集合を $U = \{u_1, u_2, \dots, u_n\}$ 、アイテムの集合を $I = \{i_1, i_2, \dots, i_m\}$ 、タグの集合を $T = \{t_1, t_2, \dots, t_r\}$ とする。このとき、三部グラフは2つの隣接行列 $A_{|U| \times |I|}$ 、 $A'_{|I| \times |T|}$ を導入することで構成される。ここで、 $A_{|U| \times |I|}$

はユーザ-アイテム二部グラフに関する隣接行列であり、 $A'_{|I| \times |T|}$ はアイテム-タグ二部グラフに関する隣接行列である。すなわち、ユーザ u_l がアイテム i_s を収集しているときに $a_{ls} = 1$ となり、そうでない場合は $a_{ls} = 0$ である。同様に、アイテム i_s にタグ t_k が割り当てられているときに $a'_{sk} = 1$ 、そうでない場合は $a'_{sk} = 0$ である。図2に3人のユーザ、5個のアイテム、4つのタグによって構成されるグラフの例を示す。ここで、各ノード間のエッジは、隣接行列において、値が1をとる要素に対応するノード間に張られていることに注意する。

§2 二部グラフ上の推薦スコアの伝搬

Zhang らの手法は、推薦スコアをユーザ-アイテム二部グラフとアイテム-タグ二部グラフ上でそれぞれ独立に伝搬させ、伝搬後に両グラフが各アイテムに与える推薦スコアを統合することによって、推薦するアイテムが決定される。本項では、ユーザ-アイテム二部グラフを例に、推薦スコアの伝搬方法について述べる。

推薦スコアの伝搬プロセスは、以下の3つのステップによって構成される:

初期状態:

ターゲットユーザが収集しているアイテムに、推薦スコアを割り当てる。

ステップ 1:

初期状態で推薦スコアを割り当てられたアイテムは、エッジで結ばれたユーザに、自身が保持する推薦スコアを均等に分配する。

ステップ 2:

ステップ 1 で推薦スコアを分配されたユーザは、自身が保持する推薦スコアを、均等にエッジで結ばれたアイテムに再分配する。

また、初期状態における推薦スコアベクトルを $\vec{f}$ 、伝搬後の推薦スコアベクトルを $\vec{f}'$ とすれば、推薦スコアの伝搬プロセスは以下のように定式化される:

$$f'_j = \sum_{l=1}^n \frac{a_{lj}}{k(u_l)} \sum_{s=1}^m \frac{a_{ls}}{k'(i_s)} f_s, \quad j = 1, 2, \dots, m.$$

ここで、 $k(u_l) = \sum_{j=1}^m a_{lj}$ はユーザ u_l によって収集されたアイテムの数であり、 $k'(i_s) = \sum_{j=1}^n a_{js}$ はアイテム i_s を収集したユーザの数である。また、ターゲットユーザを u_t とすると、ユーザ u_t における初期状態の推薦スコアベクトル $\vec{f}(u_t)$ は、以下のように定義される:

$$f_j(u_t) = a_{tj}, \quad j = 1, 2, \dots, m.$$

同様の伝搬プロセスが、アイテム-タグ二部グラフにも適用される。図2に、推薦スコアが伝搬される様子を示す。

§3 推薦スコアの統合

Zhang らの手法では、ユーザに推薦するアイテムはユーザ-アイテム二部グラフ、およびアイテム-タグ二部グラフ上の推薦スコアの伝搬結果を統合することにより決定される。ユーザ-アイテム二部グラフ、アイテム-タグ二

図2 Zhang らの手法 [Zhang 10] の概要

部グラフの伝搬後の推薦スコアベクトルをそれぞれ $\vec{f}'_u$, $\vec{f}'_t$ とすれば, 統合後の推薦スコアベクトル $\vec{f}^*$ は $\vec{f}'_u$ および $\vec{f}'_t$ の線形補間で表現される:

$$\vec{f}^* = \alpha \vec{f}'_u + (1 - \alpha) \vec{f}'_t.$$

ここで, $\alpha \in [0, 1]$ は $\vec{f}'_u$ と $\vec{f}'_t$ が推薦に寄与する度合いを調整するパラメータである. また, $\vec{f}'_u$ と $\vec{f}'_t$ はターゲットユーザごとに異なることに注意する. 最終的に, $\vec{f}^*$ のうち高い推薦スコアを持つアイテムが推薦される.

4. 提案手法

図3に提案手法の全体構成を示す. 提案手法では, Web ページ推薦システムの推薦根拠が与えられることが前提となっている. まず, Web ページ推薦システムがユーザに対して推薦するページを決定する. このとき, Web ページ推薦システム内部には, 各推薦ページの推薦根拠が保持されている. 推薦ページの決定後, システムが保持している推薦根拠と個々の推薦ページを入力とし, 単語重要度を決定する. 推薦根拠を用いて個々の推薦ページ内に含まれる単語の重要度を決定することにより, 単語重要度にユーザの嗜好を反映させることが可能である. 最終的に, 決定した単語重要度を文書要約モデルに適用することによって, ユーザ指向型スニペットを生成する.

4.1 Web ページ推薦システムにおける推薦根拠の定式化

本節では, 推薦根拠をシステム上で取り扱うための定式化を行う. 定式化に先立ち, 導入として推薦手法の一般的な枠組みについて概観する.

はじめに, 本稿では便宜上, ユーザやアイテム, タグなどの, 推薦を生成するのに利用する個々のデータを総じてリソースと呼ぶ. 前章の代表的な推薦システムのアルゴリズムで概観したように, 一般的に, Web ページ推薦システムはリソース同士の共起情報を活用し, 2つの異なるリソース間の関係の強さに基づいて, 個々のリソースのスコアを計算する. ここで, リソースが共起するとは, ある2つのリソース間に何らかの関係が与えられることである. たとえば, ユーザがなんらかのアイテムをお気に入り登録した場合, ユーザとアイテムが共起しているとみなし, あるいは, アイテムになんらかのタグが

図3 提案手法の全体構成 (赤色部分が本研究による提案部分)

付与された場合は, アイテムとタグが共起しているとみなす. そして, 共起関係から算出された個々のリソースのスコアを集約することで, ユーザに推薦するアイテムが決定される. たとえば, ユーザ間型メモリベース法では, ユーザ間の類似度がリソース間の関係の強さであり, ユーザ間の類似度によって重み付けされた評価値が, リソースのスコアに相当する. また, アイテム間型メモリベース法では, アイテム間の類似度がリソース間の関係の強さであり, アイテム間の類似度によって重み付けされた評価値が, リソースのスコアに相当する.

推薦根拠を定式化するために, 我々はユーザにアイテムを推薦する際に寄与したリソースとスコアに着目する. 上記の推薦における一般的な戦略に従うと, Web ページ推薦システムは, あるアイテムの推薦に寄与したリソースとそのスコアの対を保持していることになる. そこで, 提案手法では, あるアイテムの推薦に寄与したすべてのリソース-スコア対を推薦根拠とみなす. 提案手法における推薦根拠の考え方は, リソース-スコア対におけるスコアを, 対となるリソースが推薦に寄与した度合いと考え, この度合いを活用して単語の重要度を決定する (4.2 節で詳述する). 定式化すると, アイテム i_n がターゲットユーザ u_t に推薦されたときの推薦根拠 $Rsn(u_t, i_n)$ は, 以下のようになる:

$$Rsn(u_t, i_n) = \{(h(u_t, r_l, i_n), r_l) | r_l \in \mathbf{R}\}. \quad (1)$$

ここで, $h(u_t, r_l, i_n)$ はユーザ u_t の推薦を生成する際にリソース r_l がアイテム i_n に与えたスコアを表し, $\mathbf{R}$ は Web ページ推薦システムが推薦を生成する際に利用したリソースの集合である. $\mathbf{R}$ および $h(u_t, r_l, i_n)$ の値域は, 推薦に利用する推薦システムに依存する. たとえば, ユーザ間型メモリベース法 [Adomavicius 05] においては, 推薦されるアイテムは, ターゲットユーザとその他のユーザとの類似度に基づいたスコアを集約することにより決定されるため, リソースはユーザとなる. また, アイテ

ム間型メモリベース法 [Sarwar 01] においては、ユーザが過去に評価したアイテムと推薦されたアイテムの類似度に基づくスコアが集約されることで推薦が行われているため、リソースはアイテムとみなせる。また、このような推薦根拠の定式化は、0, 1 の二値で表される評価値データに限らず、5 段階の評価値データなど、様々なスケールの評価値データに適用できる。このように定義された Rsn は、ユーザ指向型スニペットを生成する際の単語の重要度を決定するために用いられる。

4.2 ユーザ指向型スニペット生成

§1 文書要約モデル

ユーザ指向型スニペットの生成には、文抽出による文書要約モデルを用いる。ここで、ユーザ指向型スニペットの生成には、様々な文書要約モデルが利用可能であることに注意する。本稿では、Takamura ら [Takamura 09] が提案した最大被覆問題の変種による文書要約モデルを用いる。Takamura らのモデルは文書要約を最大被覆問題として捉え、冗長性を排除することで、重要度の高い単語をより多く被覆した要約を生成するものである。我々の仮定は、ユーザが興味を持つ単語をより多く被覆したスニペットは、ユーザの適合性判定に寄与するというものである。しかしながら、冗長性の排除は今回のタスクにおいて有効に働かなかったことを 6 章において述べる。ユーザ指向型スニペットを生成するために、本研究では、ターゲットユーザ u_t の単語 e_j に対する重要度 w_{tj} をモデルに組み込むことを考える。従って、本研究が扱う文書要約モデルは、以下の整数計画問題として定式化される：

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

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、ユーザが興味を持つ単語をより多く被覆したスニペットが生成可能となる。

本研究が扱うタスクは、単一文書要約とみなせることに注意する。単一文書要約においては、文の文書中での出現位置に関する情報が有効に働くことが知られている [奥村 05]。特に、文書の先頭に近い文ほど重要な情報が含まれている場合が多いことに着目し、以下のように Takamura らのモデルの目的関数を拡張する：

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

ここで $\text{pos}(s_i)$ は、文 s_i が文書中に出現する位置を表し、 s_i が先頭から何文目にあるかによって決定される。 μ は文の位置情報を考慮する度合いを制御するパラメータであり、 μ が大きいほど提案手法は先頭に近い文を重要とみなす。また、制約条件は同様である。

§2 単語重要度 w_{tj} の決定

単語重要度 w_{tj} を決定する際には、4.1 節で定義した推薦根拠を活用する。単語重要度 w_{tj} の決定における基本的な考え方は、推薦ページに対して高いスコアを与えたりリソース、すなわち、推薦に寄与した度合いが大きいリソースと共起するページの特徴語は重要だとみなすことである。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 i_n がユーザ u_t に与えられ、その推薦根拠が $Rsn(u_t, i_n)$ であるとき、以下のような式に基づき w_{tj} を決定する：

$$w_{tj} = \frac{1}{Z} f(u_t, i_n, e_j),$$

$$f(u_t, i_n, e_j) = \sum_{(h,r) \in Rsn(u_t, i_n)} h \times TF(r, e_j) \times IDF(e_j).$$

ここで Z は正規化定数であり、 $Z = \sum_j f(u_t, i_n, e_j)$ である。また、 $TF(r, e_j)$ はリソース r と共起した Web ページ集合における単語 e_j の頻度であり、 $IDF(e_j)$ はシステムが取り扱うすべての Web ページにおける逆文書頻度である。これにより、単語の重要度にユーザの嗜好を反映させることが可能となる。

5. 実 験

本章では、提案手法の評価実験について述べる。まず、実験に用いる Web ページ推薦システムについて述べる。このとき、実験で用いる Web ページ推薦システムの推薦根拠についても言及する。その後、実験で利用するデータセット、正解データの作成方法について詳述し、具体的な実験設定および実験結果を述べる。

5.1 Web ページ推薦システムおよび推薦根拠

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を用いる。具体

的には、3.2 節で述べた、CTS に基づく Zhang らの手法 [Zhang 10] を用いる。

Zhang らの手法を振り返ると、推薦するアイテムは、ステップ 2 において、ユーザとタグによって与えられるスコアを統合することによって決定されている。従って、式 (1) に従い、Zhang らの手法の推薦根拠は以下のように定義される：

$$Rsn(u_t, i_n) = \{(h(u_t, r_l, i_n), r_l) | r_l \in U \cup T\}.$$

ここで、 u_t はターゲットユーザ、 i_n は推薦ページである。また、 $h(u_t, r_l, i_n)$ は $r_l \in U \cup T$ から推薦ページへ与えられるスコアである。

5.2 実験で利用するデータセット

本実験では livedoor クリップデータセット^{*6}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

表 1 に、データセットの内訳を示す。最終的に、データセットは 2,458 人のユーザ、10,501 個の Web ページ、4,027 個の異なり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。

5.3 正解データの作成

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適切なデータセットがないため、11 人のユーザの協力を得て、以下の (1)、(2)、(3) に従い正解データの作成を行った。

(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、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うように指示した。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

表 1 データセットの内訳

$ U $	$ T $	$ I $	ブックマーク数	タグ付け数
2,458	4,027	10,501	101,792	227,025

ク登録を行うように指示し、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うよう指示した。Web ページを提示する際には、ユーザの負担を軽減するために、ランダムにページを提示する方法とキーワード検索により提示する方法の 2 種類を用意した。また、提示するページはデータセットに含まれるものとした。

(2) Web ページの推薦

前述した Zhang ら [Zhang 10] 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。ここで、 α は 0.5 とし^{*7}、ユーザごとに 10 件のページを推薦した。また、ユーザに対して推薦されたページを全て読み、それぞれのページに興味があったか否かを判定するように指示した。

(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、ユーザが興味があったページに対して、自身が欲しいと考えるスニペットを生成するように指示した。興味がなかったページについては正解スニペットを生成するのが困難なため、今回の評価では正解スニペットの生成を行っていない。ユーザにスニペットを生成するように指示した際には、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 文字以下で選択するように依頼した。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、これらを正解データとした。

実験に先立ち、作成した正解データの分析を行った。具体的には、正解スニペットとして選ばれた文が、記事中のどの範囲から選ばれたかを調査した。図 4 に、記事を一定の範囲で区切り、範囲内の文のうち正解スニペットとして選ばれた文の割合を示す。図のスコアは、記事中の文を先頭から 10% ずつの範囲に区切り、個々の範囲に含まれる文のうち正解スニペットとして選ばれた文の割合を算出したものである。図 4 を見ると、記事の冒頭に近い文ほど、正解スニペットに含まれる傾向があることがわかる。特に、先頭から 20% 以内に含まれる文が正解スニペットに含まれる割合が高い。単一文書要約の研究においては、記事の冒頭に重要な情報が書かれていることが知られており [Brandow 95, Nenkova 05]、今回作成した正解データも同様の傾向を持つことがわかる。

^{*6} 現在は公開を終了している。

^{*7} [Zhang 10] において、 $\alpha = 0.5$ のときに、推薦リスト内の適合アイテム数が増加する結果が示されているため、この値に固定した。

図 4 特定の範囲内の文のうち正解文として選ばれた文の割合

5.4 実験設定

§1 前処理

データセットに含まれる Web ページに対して、Web ページの本文抽出、文分割、形態素解析を行った。本文抽出には ExtractContent^{*8}を利用し、不要な HTML タグの除去を行った。文分割は句読点、感嘆符、および BR タグを手掛かりに行った。形態素解析器は MeCab[Kudo 04]^{*9}を用いた。また、文書要約モデルを用いてスニペットを生成する際は、名詞、動詞、形容詞のみを考慮した。

§2 従来手法との比較

提案手法のうち文の位置情報を考慮しないモデルを $Proposed_{org}$ 、考慮したモデルを $Proposed_{pos}$ と表記する。実験のベースラインとして、以下の手法を用いる：

$Lead_{max}$ ：

リード文、すなわち本文を先頭から順次抽出しスニペットとする手法

$Lead_{ave}$ ：

$Lead_{max}$ において、最大長を $Proposed_{pos}$ および $Generic_{pos}$ が生成したスニペットの平均長と揃える手法

CS ：

Collaborative Summarization[Qu 09]、すなわち Qu らによって提案された PDS モデルを用いてスニペットを生成する手法

$Generic$ ：

提案手法において、単語重要度の決定に推薦根拠を用いない手法

$Generic_{pos}$ ：

$Proposed_{pos}$ と同様に、 $Generic$ に文の位置情報を組み込んだ手法

このときスニペットの最大長は 150 文字とした。 $Lead_{ave}$ においては、 $Proposed_{pos}$ 、 $Generic_{pos}$ が生成したスニペットの平均長が共に 146 文字であったため、最大長を 146 文字とした。また、 $Lead_{max,ave}$ においては、文を順次抽出した結果、生成されるスニペットが最大長以上と

なってしまった場合は、最後の一文が部分文となることを許した。このため、 $Lead_{max,ave}$ は他の手法と比較して有利な設定であることに注意する。公正な比較のために、今回は比較対象として $Lead_{ave}$ を用いるが、参考のために $Lead_{max}$ による実験結果も記載する。 $Generic$ 、 $Generic_{pos}$ においては、単語重要度の決定には TF-IDF を利用した。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。整数計画問題を解く際には CPLEX^{*10}を用いた。

5.5 ユーザ指向型スニペットの評価

評価尺度としては、ROUGE-2[Lin 04]を用いた^{*11}。ROUGE-2 は、文書要約の評価において広く用いられている自動評価指標である。ROUGE-2 の数値は、システムが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表す。特に、ROUGE-2 は単一文書要約において、人間の評価と非常に高い相関を示すことが報告されている [Lin 04]。本研究のタスクも単一文書要約の一種とみなせることより、ROUGE-2 を評価尺度として採用した。

評価の際には、ユーザ単位で leave-one-out 交差検定を行った。すなわち、ターゲットユーザ 11 人のうち、1 人のユーザが作成した正解スニペットをテスト事例とし、残りの 10 人が作成した正解データを用いてパラメータの調整を行うということを、各ユーザが 1 回ずつテスト事例となるよう 11 回繰り返した。このとき、各手法のパラメータは、ROUGE-2 スコアの平均が最大となる値に調整した。調整を行ったパラメータは、 $Proposed_{org}$ 、 $Generic$ については $\lambda(\in [0.1, 0.2, \dots, 1.0])$ 、 $Proposed_{pos}$ 、 $Generic_{pos}$ については、 λ および $\mu(\in [0.1, 0.2, \dots, 3.0])$ である。 α を変化させると、提案手法の単語重要度は変化するが、正解スニペットを作成した記事がターゲットユーザに対する推薦リストに出現しなくなる状況が生じるため、正解データ作成時と同様に $\alpha = 0.5$ で固定した。また、 CS においては personalization vector の更新に用いるパラメータ d を調整し、その他のパラメータは [Qu 09] と同じ値を用いた^{*12}。詳細は [Qu 09] を参照されたい。

5.6 実験結果

各手法の ROUGE-2 スコアの比較を表 2 に示す。表 2 中に示した数値は、ROUGE-2 スコアの平均である。位置情報を用いる $Proposed_{pos}$ 、 $Lead_{ave}$ 、 $Generic_{pos}$ 間でのスコアの差と、位置情報を用いない $Proposed_{org}$ 、 $Generic$ 、 CS 間のスコアの差に対してそれぞれ並べ替え検定を行った結果、有意差は確認できなかった。

*10 <http://www-01.ibm.com/support/docview.wss?uid=swg24036489>

*11 ROUGE-2 の計算においてはストップワードの除去は行わないようにした。すなわち、評価スクリプト実行時のオプション “-a -n 2 -x” に相当する。

*12 d 以外のパラメータに関しては、先行研究で有効に働く数値が言及されているため。

*8 <http://search.cpan.org/~tarao/HTML-ExtractContent-0.10/lib/HTML/ExtractContent.pm>

*9 MeCab の辞書には、IPA 辞書に Wikipedia のタイトルを追加した辞書を使用した。

表 2 を見ると, $Proposed_{org}$ は, $Generic$ および CS を上回っていることがわかる. これは, 推薦根拠を用いた単語重要度の決定が, ROUGE-2 スコアの向上に寄与することを示唆している. 言い換えると, $Proposed_{org}$ が出力したスニペットは, $Generic$, CS よりも適合性判定に有用な情報を多く含んでいると言える.

また, $Proposed_{pos}$ は, $Lead_{ave}$ および $Generic_{pos}$ を上回っている. ここで, $Proposed_{pos}$ および $Generic_{pos}$ は, スニペット中に文のみを含めることしかできない制約が課せられているのに対し, $Lead_{ave}$ は部分文も許容されている必要がある. $Proposed_{pos}$ はこのような制約下でありながら, $Lead_{ave}$ と同等の性能を発揮しているといえる.

文の位置情報を用いる $Proposed_{pos}$, $Lead_{ave}$, および $Generic_{pos}$ は, 位置情報を用いない手法, すなわち $Proposed_{org}$, $Generic$, CS と比較して大幅に高い ROUGE-2 スコアを獲得している. 5.3 節で示した通り, 正解スニペット中の文の多くが記事冒頭の文であるため, 位置情報を考慮する手法が冒頭文に高い重みを与えた結果, 高い ROUGE-2 スコアを獲得していると考えられる. これは, ページの冒頭に記述されている情報は, 適合性判定に有用な情報を含んでいることを示している. しかしながら, $Lead_{ave}$ のスコアは 0.277 にとどまっており, 今回の実験設定ではオラクル ROUGE-2 スコアが 1.0 であることを鑑みると, ページの冒頭の文だけでは, 適合性判定に必ずしも十分であるわけではないことも示唆される.

6. 考 察

冗長性を制御するパラメータ λ の効果を確認するため, 各手法の λ の変化に対する ROUGE-2 スコアの変化を調査した. 結果を図 5 に示す. ここで, $Proposed_{pos}$ および $Generic_{pos}$ については, それぞれパラメータ μ の値を 0.8, および 0.6 に固定した上で λ の値を変化させた. この μ の数値は, 両手法のパラメータチューニングを行った際に選択された値の中央値である. 図を見てわかるように, λ の増加に従い ROUGE-2 スコアは上昇する傾向にあることがわかる. これは今回のタスクにおいて冗長性の排除は ROUGE-2 スコアの向上に寄与しないことを示唆している. このような結果が得られた理由としては, 個々の Web ページの冗長性があまり高くないことが挙げられる. そのため, 冗長性の排除によりユーザの適合性判定に寄与しない文, すなわちユーザが興味のない情報を含む文がスニペットに含まれてしまった可能性がある.

続いて, 位置情報を考慮する手法ごとの傾向を明らかにするため, 位置情報を考慮する各手法が相対的に高い ROUGE-2 スコアとなった事例の調査を行った.

まず, $Proposed_{pos}$ に関して調査を行った. その結果, $Proposed_{pos}$ はターゲットユーザの嗜好を考慮した単語の重み付けにより, 高い ROUGE-2 スコアを獲得するス

表 2 各手法の ROUGE-2 スコアの比較

手法	ROUGE-2
$Proposed_{pos}$	0.282
$*Lead_{max}$	0.281
$Lead_{ave}$	0.277
$Generic_{pos}$	0.274
$Proposed_{org}$	0.132
$Generic$	0.120
CS	0.101

図 5 λ の変化による ROUGE-2 スコアの変化

ニペットを生成できたと考えられる. 一例として, 表 3 に $Proposed_{pos}$ が最も高い ROUGE-2 スコアを獲得した, 推薦対象記事の冒頭と正解スニペット, ターゲットユーザの嗜好情報の一部, および各手法により生成されたスニペットの一部を示す. 各手法が生成したスニペットにおいて, 正解スニペットに含まれている文には下線を引いている. 各手法のスニペットを比較すると, $Proposed_{pos}$ は「クックパッドの一番の目標は何か?」という文を抽出しているため, 高い ROUGE-2 スコアを獲得している. $Lead_{ave}$ では, この文が最大長を超えた位置に出現するため, スニペットに含めることができていない. 一方, $Generic_{pos}$ では「クックパッド」という単語に対する重みが低いため, この文を抽出できていないと考えられる. 提案手法と generic な手法による「クックパッド」の重みを比較すると, それぞれ 0.15, 0.05 であった. ターゲットユーザはクックパッドに関するページをブックマークしており, このような重み付けは提案手法がユーザの嗜好を捉えた結果であると考えられる.

続いて, $Generic_{pos}$ に関して調査を行った. 調査の結果, $Generic_{pos}$ は, 正解スニペットに含まれている単語が推薦対象記事内で頻出する場合に, 高い ROUGE-2 スコアを獲得すると思われる. その一例として, 表 4 に, $Generic_{pos}$ が最も高い ROUGE-2 スコアを獲得した例を示す. この例の推薦対象記事内には, 正解スニペッ

表 3 $Proposed_{pos}$ が高い ROUGE-2 スコアを獲得した例

ターゲットユーザの嗜好情報		
タグ	ブックマークページ	
rails プログラミング, iPhone 自然言語処理 経済, 旅行 イノベーション	「巨大サイトでも普通に Rails で動いてますけど!」クックパッド&カカクコム共催・勉強会の亀レポート iPhone 開発で役立ちそうなチュートリアルを紹介 機械学習による自然言語処理チュートリアル バンコクで「物価」の意味を知る デザイン思考の道具箱 イノベーションを生む会社のつくり方	
推薦対象記事 (記事タイトル: 月間 2.8 億 pv! クックパッドの運営ノウハウについて聞いてきた)		
記事の冒頭 *13	月間 2.8 億 pv! クックパッドの運営ノウハウについて聞いてきた 百式さんのカジュアル勉強会が、今回はクックパッドのオフィスで開催されました。テーマは「コミュニティ運営」について*1。噂のクックパッドオフィスにも潜入してきましたよ! クックパッドを立ち上げた佐野さんにお話を伺いました。進行は IDEA*IDEA のカジュアル勉強会の進め方に乗っ取った感じで。ノート PC 持ってくるのを忘れたので、以下メモ書きから印象に残っている質問&回答を... クックパッドの一番の目標は何か? 料理が楽しくなること。サイトを大きくしたい、...	
正解スニペット	クックパッドの運営ノウハウについて聞いてきた ... クックパッドを立ち上げた佐野さんにお話を伺いました。... クックパッドの一番の目標は何か? 料理が楽しくなること。サイトを大きくしたい、ユーザ数を増やしたいと思ったことはない。	
各手法が生成したスニペット		
手法	スニペット	ROUGE-2
$Lead_{ave}$	月間 2.8 億 pv! クックパッドの運営ノウハウについて聞いてきた 百式さんのカジュアル勉強会が、今回はクックパッドのオフィスで開催されました。テーマは「コミュニティ運営」について*1。噂のクックパッドオフィスにも潜入してきましたよ! <u>クックパッドを立ち上げた佐野さんにお話を伺いました。</u> 進行は ...	0.429
$Generic_{pos}$	月間 2.8 億 pv! クックパッドの運営ノウハウについて聞いてきた 百式さんのカジュアル勉強会が、今回はクックパッドのオフィスで開催されました。... 噂のクックパッドオフィスにも潜入してきましたよ! ... クックパッドを立ち上げた佐野さんにお話を伺いました。... 月間 446 万 uu、月間 2.8 億 pv。	0.429
$Proposed_{pos}$	月間 2.8 億 pv! クックパッドの運営ノウハウについて聞いてきた 百式さんのカジュアル勉強会が、今回はクックパッドのオフィスで開催されました。... 噂のクックパッドオフィスにも潜入してきましたよ! ... <u>クックパッドを立ち上げた佐野さんにお話を伺いました。... クックパッドの一番の目標は何か? 料理が楽しくなること。</u>	0.653

ト 1 文目に出現する「書籍」,「新刊」,「オンライン書店」といった単語が頻出する。そのため、 $Generic_{pos}$ は、 $Proposed_{pos}$ が抽出できていない正解スニペット 1 文目を抽出することができたと考えられる。

最後に、 $Lead_{ave}$ については、特に正解スニペットがリード文のみで構成される場合に、高い ROUGE-2 スコアを獲得すると考えられる。このような事例が、正解スニペット全 50 事例中、2 事例存在した。このような場合、 $Lead_{ave}$ 以外の手法では、単語の重み付けによってリード文以外の文を抽出してしまうため、ROUGE-2 スコアが低くなる。今回の実験データではリード文が有効に働いたが、一方で、ドメインによってはリード文が有効に働かないことが知られている [菊池 15]。今回実験で使ったブログ記事ではリード文が強い記事が多かったが、 $Proposed_{pos}$ はパラメータを調整することで位置情報が作用する強さを調整できるため、リード文が有効でないドメインに対して有効に働く可能性がある。

本研究の評価では、正解スニペットを生成したページは、ユーザが興味を持つページのみに限定されていることに注意する必要がある。現状の評価では、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悪欠している。この観点についての評価は、今後の課題とする。

*13 原文をそのまま記載しているため、誤字・脱字の訂正は行っていない。

7. ま と め

本研究では、Web ページ推薦システムにおけるユーザの適合性判定を支援するために、ユーザ指向型スニペットを生成するためのモデルを提案した。提案手法では、Web ページ推薦システムの推薦根拠を活用したモデル化を行った。具体的には、推薦根拠と関連が深い単語は重要であるとみなし、重要な単語を求めた後、文書要約モデルを適用することでスニペットの生成を行った。推薦根拠の活用により、スニペットにユーザの嗜好を反映させることが可能となる。また、提案手法は、Web ページ推薦システムの推薦根拠を、推薦に利用した個々のデータと、個々のデータが推薦に寄与した度合いを表すスコアの対として与えることができれば、様々な Web ページ推薦システムにおいて利用可能である。基本的に、推薦システムでは、個々のデータが推薦に寄与する度合いをスコアとして算出し、算出したスコアを集約することで推薦を生成する。そのため、多くの Web ページ推薦システムにおいて、このような形式で推薦根拠を与えることが可能であると考えられる。

人手で作成したデータセットを用いた評価実験により、提案手法は先行研究、および generic な文書要約モデルよりも高い ROUGE-2 スコアを獲得した。また、提案手法はスニペットに文のみを含める制約下にありながら、部分文を許容する $Lead$ 法と同等の性能を示した。

表 4 $Generic_{pos}$ が高い ROUGE-2 スコアを獲得した例

ターゲットユーザの嗜好情報		
タグ	ブックマークページ	
本, book, 書評, 作文 ruby linux, shell lifehack, 仕事術, 手帳 カラーパレット, 浮世絵	学生必読「理科系の作文技術」 x86_64 環境で Ruby から MySQL のクエリを実行するときの問題が示す根本的な問題 添削:業務で楽するための UNIX テクニック集 まずはおさらい、シェル制御構造と正規表現の基礎 年収 2000 万円の手帳術 浮世絵に学ぶ日本の伝統色と配色	
推薦対象記事（記事タイトル：なかなか見つからない本を検索したり、本を管理する WEB サービスまとめ）		
記事の冒頭	なかなか見つからない本を検索したり、本を管理する WEB サービスまとめ book, search 稀観書 (希少な古書籍) を古書肆 (古本屋) やネットオークションで探し回ったり、在庫僅少の新刊書籍をオンライン書店で検索したり。買った後に山積みになった本を何とか管理しようとして書架 (本棚) を整理したり。三度の飯より読書や書肆めぐりが好きで、活字中毒/本の虫/書痴/書狂/愛書家/書物崇拜狂/Bibliophilia/ビブロフィリア/bibliophile/ビブリオファイルの自分が役に立ったウェブサービスをまとめてみた。「Amazon ほしい物リスト」*1 しか知らなかった ...	
正解スニペット	稀観書 (希少な古書籍) を古書肆 (古本屋) やネットオークションで探し回ったり、在庫僅少の新刊書籍をオンライン書店で検索したり。 ... S T E P 1 まずは読みたい本のテーマから類書を探す (または Q&A サイトで質問して探す) ... S T E P 4 新刊で入手できない希少本・絶版本は古本屋に出回っていないか検索	
各手法が生成したスニペット		
手法	スニペット	ROUGE-2
$Lead_{ave}$	なかなか見つからない本を検索したり、本を管理する WEB サービスまとめ book, search 稀観書 (希少な古書籍) を古書肆 (古本屋) やネットオークションで探し回ったり、在庫僅少の新刊書籍をオンライン書店で検索したり。買った後に山積みになった本を何とか管理しようとして書架 (本棚) を整理したり。	0.403
$Generic_{pos}$	なかなか見つからない本を検索したり、本を管理する WEB サービスまとめ ... 稀観書 (希少な古書籍) を古書肆 (古本屋) やネットオークションで探し回ったり、在庫僅少の新刊書籍をオンライン書店で検索したり。 ... S T E P 1 まずは読みたい本のテーマから類書を探す (または Q&A サイトで類書を探す (または Q&A サイトで質問して探す) ... 古本販売サイト一括検索	0.736
$Proposed_{pos}$	なかなか見つからない本を検索したり、本を管理する WEB サービスまとめ book, search ... 買った後に山積みになった本を何とか管理しようとして書架 (本棚) を整理したり。 ... S T E P 1 まずは読みたい本のテーマから類書を探す (または Q&A サイトで質問して探す) ... あの本のタイトルが知りたい：復刊ドットコム	0.375

今後の課題として、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と同様に、より実際の運用に近い形で評価することが挙げられる [Tombros 98]。加えて、協調フィルタリングなどの、その他の推薦手法に対しても提案手法を適用することも検討している。

謝 辞

本研究は JSPS 科研費 26280080 の助成を受けたものである。

◇ 参 考 文 献 ◇

- [Adomavicius 05] Adomavicius, G. and Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, No. 6, pp. 734–749 (2005)
- [Brandow 95] Brandow, R., Mitze, K., and Rau, L. F.: Automatic condensation of electronic publications by sentence selection, *Information Processing & Management*, Vol. 31, No. 5, pp. 675–685 (1995)
- [Díaz 07] Díaz, A. and Gervás, P.: User-model based personalized summarization, *Information Processing & Management*, Vol. 43, No. 6, pp. 1715–1734 (2007)
- [Golder 06] Golder, S. A. and Huberman, B. A.: Usage patterns of collaborative tagging systems, *Journal of Information Science*, Vol. 32, No. 2, pp. 198–208 (2006)
- [Hotho 06] Hotho, A., Jäschke, R., Schmitz, C., Stumme, G., and Althoff, K.-D.: FolkRank: A ranking algorithm for folksonomies, in *Proceedings of Workshop on Information Retrieval 2006 (FGIR)*, Vol. 1, pp. 111–114 (2006)
- [菊池 15] 菊池 悠太, 平尾 努, 高村 大也, 奥村 学, 永田 昌明: 入れ子依存木の刈り込みによる単一文書要約手法, 自然言語処理, Vol. 22, No. 3, pp. 197–217 (2015)
- [Kudo 04] Kudo, T., Yamamoto, K., and Matsumoto, Y.: Applying conditional random fields to Japanese morphological analysis, in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 230–237 (2004)
- [Lin 04] Lin, C.-Y.: ROUGE: A package for automatic evaluation of summaries, in *Proceedings of the Workshop on Text Summarization Branches Out*, pp. 74–81 (2004)
- [Nenkova 05] Nenkova, A.: Automatic text summarization of newswire: Lessons learned from the document understanding conference, in *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI)*, Vol. 5, pp. 1436–1441 (2005)
- [奥村 05] 奥村 学, 難波 英嗣: テキスト自動要約, オーム社 (2005)
- [Park 12] Park, D. H., Kim, H. K., Choi, I. Y., and Kim, J. K.: A literature review and classification of recommender systems research, *Expert Systems with Applications*, Vol. 39, No. 11, pp. 10059–10072 (2012)
- [Qu 09] Qu, Y. and Chen, Q.: Collaborative summarization: When collaborative filtering meets document summarization, in *Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation (PACLIC)*, pp. 474–483 (2009)
- [Sarwar 01] Sarwar, B., Karypis, G., Konstan, J., and Riedl, J.: Item-based collaborative filtering recommendation algorithms, in *Proceedings of the 10th International Conference on World Wide Web (WWW)*, pp. 285–295 (2001)
- [Takamura 09] Takamura, H. and Okumura, M.: Text summarization model based on maximum coverage problem and its variant, in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pp. 781–789 (2009)
- [Tombros 98] Tombros, A. and Sanderson, M.: Advantages of query biased summaries in information retrieval, in *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and De-*

velopment in Information Retrieval (SIGIR), pp. 2–10 (1998)

- [Turpin 09] Turpin, A., Scholer, F., Jarvelin, K., Wu, M., and Culpepper, J. S.: Including summaries in system evaluation, in *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 508–515 (2009)
- [Wan 07] Wan, X. and Xiao, J.: Towards a unified approach based on affinity graph to various multi-document summarizations, in *Proceedings of the 11th European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, pp. 297–308 (2007)
- [Wang 07] Wang, C., Jing, F., Zhang, L., and Zhang, H.-J.: Learning query-biased web page summarization, in *Proceedings of the 16th ACM Conference on Information and Knowledge Management (CIKM)*, pp. 555–562 (2007)
- [Wetzker 09] Wetzker, R., Umbrath, W., and Said, A.: A hybrid approach to item recommendation in folksonomies, in *Proceedings of the WSDM'09 Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR)*, pp. 25–29 (2009)
- [White 02] White, R. W., Ruthven, I., and Jose, J. M.: Finding relevant documents using top ranking sentences: an evaluation of two alternative schemes, in *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 57–64 (2002)
- [White 03] White, R. W., Jose, J. M., and Ruthven, I.: A task-oriented study on the influencing effects of query-biased summarisation in web searching, *Information Processing & Management*, Vol. 39, No. 5, pp. 707–733 (2003)
- [Xu 06] Xu, Y., Zhang, L., and Liu, W.: Cubic analysis of social bookmarking for personalized recommendation, in *Frontiers of WWW Research and Development-APWeb 2006*, pp. 733–738, Springer (2006)
- [Yan 11] Yan, R., Nie, J.-Y., and Li, X.: Summarize what you are interested in: An optimization framework for interactive personalized summarization, in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1342–1351 (2011)
- [Zhang 03] Zhang, H., Ma, Z. C. W.-y., and Cai, Q.: A study for documents summarization based on personal annotation, in *Proceedings of the HLT-NAACL 03 on Text Summarization Workshop - Volume 5 (HLT-NAACL-DUC)*, pp. 41–48 (2003)
- [Zhang 10] Zhang, Z.-K., Zhou, T., and Zhang, Y.-C.: Personalized recommendation via integrated diffusion on user-item-tag tripartite graphs, *Physica A: Statistical Mechanics and its Applications*, Vol. 389, No. 1, pp. 179–186 (2010)

〔担当委員：風間 一洋〕

2016 年 4 月 2 日 受理

著者紹介

渡邊 亮彦

2012 年群馬工業高等専門学校専攻科生産システム工学専攻修了。2014 年東京工業大学大学院総合理工学研究科博士前期課程修了。同年、同研究科博士後期課程に進学。

笹野 遼平(正会員)

2009 年東京大学大学院情報理工学系研究科博士課程修了。博士(情報理工学)。京都大学大学院情報学研究科特定研究員を経て 2010 年より東京工業大学助教。自然言語処理、特に照応解析、述語項構造解析の研究に従事。言語処理学会、情報処理学会、ACL 各会員。

高村 大也(正会員)

1997 年東京大学工学部計数工学科卒業。2000 年同大学院工学系研究科計数工学専攻修了(1999 年はオーストリアウィーン工科大学にて研究)。2003 年奈良先端科学技術大学院大学情報科学研究科博士課程修了。博士(工学)。2003 年より東京工業大学助手、のち助教。2010 年より同准教授。計算言語学、自然言語処理を専門とし、特に機械学習の応用に興味を持つ。情報処理学会、言語処理学会、ACL 各会員。

奥村 敦(正会員)

1989 年東京工業大学大学院理工学研究科情報工学専攻博士課程修了。同年、東京工業大学工学部情報工学科助手。1992 年北陸先端科学技術大学院大学情報科学研究科助教授。2000 年東京工業大学精密工学研究所助教授。2009 年同教授、現在に至る。工学博士。自然言語処理、自動テキスト要約、語学学習支援、テキストマイニングに関する研究に従事。情報処理学会、電子情報通信学会、AAAI、言語処理学会、ACL、認知科学会、計量言語学会各会員。