

# モバイル環境における Web 検索結果の 柔軟な表示法

目木 信太郎 †

† 岡山県立大学情報工学部

多くの場合サーチエンジンは検索結果として URL のほかに 3 行程度のテキスト (以下スニペット<sup>3)</sup> と呼ぶ) を返してくることが多い。必要とする情報が簡単なデータの場合、ユーザはこれらの情報をスニペットを読むだけで得られることも多い。モバイル環境では Web ページを閲覧するのに長い時間が掛かるので、そのような環境ではこのことは非常に重要である。しかしながら、キーワードがお互いに離れて出現する場合にはスニペットは短いいくつかのテキストの集合となるので、ユーザは必要な情報をスニペットから得られないことも多い。そのような場合、サーチエンジンは必ずしも 3 行のスニペットを返すべきではなく、必要に応じて 3 行より長いスニペットを返すべきである。検索結果のこの柔軟な表現方法は、モバイル環境においてユーザが検索結果を順次閲覧することなく必要な情報を得ることを可能にすることができる。

## A Flexible Representation Method of Web Search Results in Mobile Environments

Shintaro MEKI †

† Faculty of Computer Science & System Engineering, Okayama Prefectural University

In many cases search engines return as search results about three lines text including keywords (called snippet<sup>3)</sup>) besides URLs. If required information is simple data, users often get those information only by reading the snippets. It's very important when users are in mobile environments, because it takes long time to browse Web pages in such environments. However in the case where keywords appear apart from each other, the snippet becomes into several short pieces of text, and thus users often cannot get required information from the snippet. Then in such cases, search engines should not always return only three lines snippets, and should do longer snippets than three lines as required. This flexible representation method of search results often makes users get required information without navigating search results in mobile environments.

### 1 まえがき

インターネットは現在の情報化社会の情報インフラとなった。インターネットでは情報は Web サーバ上に存在することが多い。しかしながら、インターネットには多くの Web サーバが存在し、それぞれの Web サーバも多くの Web ページを持っている。ユーザはどの Web サーバが知りたい情報を持っているか、すなわち位置情報 (URL) を知っていなければいけない。さもなければユーザはインターネット上の情報を利用することができない。

それゆえ多くのユーザはサーチエンジンを使っ

て URL を検索する。サーチエンジンはあらかじめ作ってある索引を使って URL を検索し、多くの場合 URL のほかに約 3 行のテキスト (スニペットと呼ばれる) を URL とともに返す。ユーザはこれらのスニペットを考慮して、返された URL の中からいくつかを選択し、どの Web ページを実際に閲覧するかを決定するのである。ユーザは実際にこれらの Web ページを閲覧することによって必要とする情報を得るのである。しかしながら、必要とする情報が日付や場所や名前などといった簡単なデータの場合、スニペットがこれらの情報を含んでいることが多く、それゆえ返されたスニペット

を読むだけで必要とする情報を得ることができる。

図1を参照して頂きたい。これらの検索結果はGoogle<sup>2)</sup>を使って、「マイクロソフト」と「チェアマン」をキーワードとして検索を行った結果である。この場合ユーザは実際のWebページを閲覧する必要はない。なぜならばスニペットを読むだけでマイクロソフトのチェアマンがBill Gatesであることを理解することができるからである。

Webページには非常に重いものも存在し、そのようなWebページを閲覧するには長い時間が掛かる。モバイル環境ではブロードバンド環境よりもWebページを閲覧するのに長い時間が掛かるから、モバイル環境にいるユーザにとってはこのことは非常に深刻な問題である。それゆえスニペットを読むだけで必要とする情報を得ることができるのであれば理想的である。このアイデアはキーワードがWebページ上でお互いに近くに現れるときだけうまく働き、キーワードがお互いに離れて現れるときにはあまりうまく働かない。なぜならば、そのような場合にはスニペットは多くの短いテキストの断片になるからである。したがって必要とする情報を含んでいる確率は低くなる。本稿では、モバイル環境においてユーザが必要とする情報を実際にWebページを閲覧することなく得ることができるようにするWeb検索結果の表現方法を述べる。

本稿は以下のように構成されている。次の章では基本的なアイデアについて述べる。第3章では実験結果と考察について述べる。最後の章は本稿の結論である。

## 2 基本的なアイデア

前の章ではユーザは実際のWebページからではなくスニペットから必要な情報を得られることがあるということ述べた。このことは特にモバイル環境においてWebページを順次閲覧する時間を節約してくれる。しかしあまりうまく働かないこともある。

この章ではわれわれは検索結果を表現する方法を少し変更することを提案する。多くのサーチエンジンはURLと一緒にたった3行のスニペットを返す。そこでサーチエンジンに3行より長いスニペットを返すことを許すようにするのである。キーワードがお互いに接近してWebページ上に現れるときは、サーチエンジンが3行だけのスニペットを返すのは適切である。しかしながら、もしサーチエ

ンジンが3行より長いスニペットを返すのならば、スニペットが必要な情報を含んでいる確率はより高くなる。この検索結果の柔軟な表現方法は、モバイル環境においてユーザが検索結果のWebページを順次閲覧する時間を節約してくれる。

スニペットは必ずしも常にすべてのキーワードを含んでいるわけではない。しかしながら、スニペットが長くなるのにしたがって、スニペットが必要な情報を含んでいる確率は高くなる。それゆえ長いスニペットは必要な情報をおそらく含んでいるであろうからより望ましい。そしてモバイル環境においては検索結果を順次閲覧する時間を節約してくれる。次の章では、必要な情報を含むのに十分長く、かつ読むのに十分短いようなスニペットの望ましい長さを調べる。

なお、本稿では検索結果の表示法を提案しているのであって、新しい検索アルゴリズムを提案しているのではない。だから、サーチエンジンは自分で新たに作成してもよいし、既存のサーチエンジンの検索結果を利用してよい。

また本稿では、スニペットが3行であることにこだわらない、という意味で柔軟な検索結果の表示法という言い方をしている。

## 3 実験結果

普通ならばプロトタイプを作成して実験を行うところであるが、我々はそのようなことはしない。なぜならば、我々のアプローチは直接的であるので、例えばプロトタイプシステムを作成して実験を行っても、スニペットを長くすればするほど必要な情報が含まれる確率は高くなるという自明な結果しか得られないからである。また、インターネット上には様々なWebページが存在するので、例えばブログとそれ以外のWebページを一緒に対象にして実験を行っても、意味のある結果が得られるとは考えにくいからである。そこでプロトタイプを使った実験は行わずに、手作業で実験を行い、それをもって実験結果とする。

図2はGoogleを使ってキーワードを「タウリン」と「効能」として通常の実験を行った結果である。紙面の制約から、上位5件の検索結果だけを用いた。各検索結果には約3行のテキストのスニペットが含まれているが、それらは短すぎるのでそれらの意味を理解するのは困難である。図3はGoogleを使って同じキーワードで、我々のアイ

1. 【楽天市場】 マイクロソフト Halo3 リミテッドエディション 【Xbox360用 ...  
 マイクロソフト Halo3 リミテッドエディション 【Xbox360用】 【税込】 9UE-00010 ヘイロー  
 3 リミテッド [9UE00010 ヘイロー 3 リミテッド] 【デジタル 0702】 商品番号, 4988648505889-  
 52-3916. メーカー希望小売価格, 7140 円 (税込). 販売価格, 6380 円 (税込) 送料 ...  
[item.rakuten.co.jp/jism/4988648505889-52-3916-n/  
 %E6%98%AF%E9%80%99%E7%89%87%E5%97%8E~%E5%A6%82%E6%9E%9C%E...](http://item.rakuten.co.jp/jism/4988648505889-52-3916-n/%E6%98%AF%E9%80%99%E7%89%87%E5%97%8E~%E5%A6%82%E6%9E%9C%E...)
2. eetimes.jp マイクロソフト社、「WinHEC」で「Windows Home Server」の ...  
 米 Microsoft (マイクロソフト) 社は 2007 年 5 月 15 日 (米国時間)、米国のカリフォルニア  
 州ロサンゼルスで開催されたハード ... Microsoft 社のチェアマンである Bill Gates 氏は  
 WinHEC の基調講演の中で、「Windows Vista は、パソコン周辺機器を大きく変えつつ ...  
[www.eetimes.jp/contents/200705/18865\\_1\\_20070520163200.cfm](http://www.eetimes.jp/contents/200705/18865_1_20070520163200.cfm)
3. マイクロソフトカップ 2006 - MSN スポーツ  
 朝日新聞社は 55 年間、全国社会人大会を主催者の一つとして応援させていただき、3 年前か  
 らマイクロソフトカップを一生懸命応援させていただいています。新聞読者の一人でも多く  
 にこの大会の意義深さを ... 中締めあいさつ ■ トップリーグ 真下昇チェアマン ...  
[inews.sports.jp.msn.com/rugby/mscup2006/pre/vip01.html](http://inews.sports.jp.msn.com/rugby/mscup2006/pre/vip01.html)
4. マイクロソフトカップ 2006 - MSN スポーツ  
 本件に関し、ジャパンラグビー トップリーグチェアマン 真下昇 (ましも・のぼる) のコメ  
 ントを以下の通り、お知らせいたします。 ... このページ上に表示される記事、あるいはリ  
 ンク先の記事内容は MSN およびマイクロソフトの見解を反映するものでは ...  
[inews.sports.jp.msn.com/rugby/mscup2006/news/20060102-16.html](http://inews.sports.jp.msn.com/rugby/mscup2006/news/20060102-16.html)
5. 米マイクロソフト、Xbox を公開 — ネット — マイコミジャーナル  
 米ラスベガスで開催されている 2001 International CES で、米マイクロソフトのチェアマン  
 &チーフソフトウェアアーキテクトのビル・ゲイツ氏は基調講演を行い、その講演において  
 新しいゲームマシン「Xbox」を公開した。この Xbox は Intel 製の 733MHz の CPU、 ...  
[journal.mycom.co.jp/news/2001/01/07/12.html](http://journal.mycom.co.jp/news/2001/01/07/12.html)

Fig. 1 簡単な質問の検索結果

### 1. かすみ日記 Blog: タウリンの効能

タウリンの効能. このところ慢性疲労がとれない。カニを食っても焼肉食っても、睡眠導入剤を飲んでも次の日の「スキッと爽快」って日がない(-;-) マルチビタミンにさらにビタミンB剤を飲んでも効かず、リボDを会社の食堂でイッキ飲みしても効かず、...

[tajima-marine-labo.cocolog-nifty.com/kasumi/2006/04/post\\_6c35.html](http://tajima-marine-labo.cocolog-nifty.com/kasumi/2006/04/post_6c35.html)

### 2. タウリン 効果と効能

タウリンの効果と効能について調べてみました。タウリンについて解説しているおすすめサイトをピックアップしています。タウリン情報収集に！

[taurine.nabigat1.net/](http://taurine.nabigat1.net/)

### 3. タウリンの効果効能:サプリメント 大辞典

タウリンの効果効能. タウリンには、肝臓でコレステロールから生合成される胆汁酸の分泌を促す作用があり、肝細胞の再生を促進します。血液中のコレステロールや中性脂肪を減少させて血液をサラサラにする働きもあり、動脈硬化の予防に役買っています ...

[www.kanemouke55.com/archives/29912390.html](http://www.kanemouke55.com/archives/29912390.html)

### 4. タウリン タウリンとは、タウリンの効能、タウリンの効果、タウリンの ...

「タウリン」に関する健康用語。タウリンとは、タウリンの効能、タウリンの効果、タウリンの栄養.

[www.karadakara.com/dict/keyword/%E3%82%BF%E3%82%A6%E3%83%AA%E3%83%B3.html](http://www.karadakara.com/dict/keyword/%E3%82%BF%E3%82%A6%E3%83%AA%E3%83%B3.html)

### 5. タウリンとは？イカ好きには朗報です。お酒のお供にタウリン！?

疲れに効くだけではない、さまざまなタウリン効能をご紹介します。滋養強壮ドリンクなどに入っているのも、体内組織には無いものと勘違いしてしまいそうですが、実は、体を作っている蛋白質であるアミノ酸の一種なのです。体内には体重の約0.1%ほど ...

[www.yamaguchi.net/archives/002416.html](http://www.yamaguchi.net/archives/002416.html)

Fig. 2 検索結果

1. かすみ日記 Blog: タウリンの効能

適応不能

[tajima-marine-labo.cocolog-nifty.com/kasumi/2006/04/post\\_6c35.html](http://tajima-marine-labo.cocolog-nifty.com/kasumi/2006/04/post_6c35.html)

2. タウリン 効果と効能

タウリン 効果と効能

まずは、タウリンって一体なに？となるべくたくさんの方のことを教えてくれるサイトさんはないかなあ  
と探してみました。難しすぎて専門家向けだなあ～というページか、情報量的にも足りないという  
ページが多い中、おすすめ度★★★★★のサイトさんがありました！

[taurine.nabigat1.net/](http://taurine.nabigat1.net/)

3. タウリンの効果効能:サプリメント 大辞典

タウリンの効果効能

タウリンには、肝臓でコレステロールから生合成される胆汁酸の分泌を促す作用があり、肝細胞の再  
生を促進します。

血液中のコレステロールや中性脂肪を減少させて血液をサラサラにする働きもあり、動脈硬化の予防  
に一役買っています。

また、心筋の収縮力を高めて心臓機能を強化する、交感神経を抑制して血圧を正常に保つ、インスリ  
ンの分泌を盛んにして血糖値を下げるといった作用があり、心臓病、高血圧、糖尿病などの予防に効果  
を発揮します。

[www.kanemouke55.com/archives/29912390.html](http://www.kanemouke55.com/archives/29912390.html)

4. タウリン タウリンとは、タウリンの効能、タウリンの効果、タウリンの ...

タウリンとは、タウリンの効能、タウリンの効果、タウリンの栄養

タウリンとは、人体に最も多く含まれているアミノ酸で、体のあちこちに存在する重要なアミノ酸の  
ひとつ。

タウリンには、コレステロールを減少させたり、心臓や自律神経の働きを良くする、疲労回復や目の  
健康維持などにも効果があるとされている。

タウリンは自然界のものでは貝や牡蠣、タコなどの魚介類に含まれている。最近では栄養ドリンクに  
もタウリンが多量に含まれていることが多い。

[www.karadakara.com/dict/keyword/%E3%82%BF%E3%82%A6%E3%83%AA%E3%83%B3.html](http://www.karadakara.com/dict/keyword/%E3%82%BF%E3%82%A6%E3%83%AA%E3%83%B3.html)

5. タウリンとは？イカ好きには朗報です。お酒のお供にタウリン！?

最近なんとなく身体がだるい、疲れがとれない…そんなあなた! そして年末年始の忘年会・新年会・ク  
リスマスパーティーでアルコールの量が増え続けるあなたにオススメ!! 疲れに効くだけではない、さ  
まざまなタウリン効果をご紹介します。

滋養強壮ドリンクなどに入っているので、体内組織には無いものと勘違いしてしまいがちですが、実  
は、体を作っている蛋白質であるアミノ酸の一種なのです。体内には体重の約0.1%ほどありますから、  
体重70kgなら70g持っていることになります。

[www.yamaguchi.net/archives/002416.html](http://www.yamaguchi.net/archives/002416.html)

Fig. 3 想像上の検索結果

デアに基づいて検索を行った想像上の検索結果である。

一番目の検索結果は特殊な場合である。Googleの検索結果では1番目になっているが、実際にWebページにアクセスしようとする、「Webページが見つかりません」となる。キャッシュでも同じことが起きるのだが、その理由はよくわからない。

3番目と4番目の検索結果について検討を行った結果、キーワードを含む行とそれに続く段落を検索結果とすればよいのではないかという結論が得られた。ここで言う段落とは、空行で区切られた一連のテキスト、またはフォントが変わって段落の切れ目と考えられるもので区切られた一連のテキストのことである。これはキーワードが小見出しとして使われている場合が多いということである。段落は明確な意味を持っているので、ユーザはその意味を理解することができると考えられる。もちろんスニペットを長くすればするほどその意味ははつきりしてくるが、あまりにも長いスニペットは特にモバイル環境では読むのに長い時間が掛る。段落は明確な意味を持つには十分長く、それを読むには十分短いと考えられる。そこで、我々はキーワードを含む行とそれに続く段落を検索結果とすればよいのではないかと考えた。インターネット上には様々なWebページが存在するので、5行や10行といった行数で決めるよりも、段落のような論理的な切れ目でスニペットを決める方がより適切であると考えられる。

2番目と5番目の検索結果はうまくいかなかった例である。3番目と4番目の検索結果と同様にキーワードを含む行とそれに続く段落を検索結果とすると、うまく意味のあるスニペットを得ることができなかった。これはキーワードと知りたい情報について書かれている部分が離れているからであると考えられる。この場合にもスニペットをさらに長くすれば、知りたい情報について書かれている部分が検索結果に含まれるようになるが、そうするとスニペットを読むのに時間が掛るようになるので、我々はそういうことはせずうまくいかない例としてあきらめることにした。

こう結論づけると、それでは不十分ではないかと考える読者もいるかも知れない。しかし、それで十分なのである。もう一度図1を参照して頂きたい。この検索結果には、ラグビーのトップリー

グのチェアマンについて書かれたWebページも含まれている。これは、マイクロソフトがラグビーのトップリーグのスポンサーになっているから検索結果に含まれているのであるが、5つの検索結果をよく検討すればこれらがノイズであって、本当に知りたいマイクロソフトのチェアマンはBill Gatesだということは理解できる。すなわち、すべての検索結果が正しい検索結果となっていないのもよいのである。

ユーザは3番目と4番目の例において、想像上の検索結果のスニペットの意味を理解することができる。ユーザがモバイル環境において、検索結果を実際に順次閲覧することなく必要な情報を得ることは可能である。

#### 4 結論

本稿では我々はスニペットを長くすることを提案し、手作業による実験結果を示した。実験結果から長いスニペットはその意味を理解ことができ、ユーザはモバイル環境においてそれらから必要な情報を得ることができ、検索結果を順次閲覧する時間を省くことができることがわかる。

我々はたった一つの実験結果を示しただけだが、提案した手法はスニペットを長くするというものであり、それゆえスニペットの情報は損なわれることはない。スニペットの持つ意味はおそらく豊かになるであろうし、少なくなることは絶対にない。したがってどんな場合にもユーザは普通の検索結果よりも多くの情報を提案した手法による実験結果から得ることができると考えられる。

本稿の最後に我々のアプローチの弱点について述べる。スニペットが長くなるとすべての検索結果をユーザが見るのに時間が掛るようになる。しかし、ユーザはブログのような長いWebページを見るのに慣れてきている。また多くのユーザはWebページを閲覧するのにスクロールマウスを使うので、我々はそれは重要な問題ではないと考えた。さらに、携帯電話のような端末でWebページを閲覧する場合、横方向のスクロールはないので、スクロールキーを押しっ放しにするだけで検索結果を見ることができる。だから、ノートパソコンの場合でも、携帯電話のような端末の場合でも、検索結果が多少長くなっても深刻な問題は発生しないと考えられる。モバイル環境においては、Webページを実際に閲覧するよりも、それらをスクロール

することによって閲覧する方がより簡単である。

#### 参考文献

- 1) Adobe: <http://www.adobe.com/>.
- 2) Google: <http://www.google.com/>.
- 3) 高見真也, 田中克己: ウェブページに対する定量的評価の視覚化による情報検索支援, DB-Web2006, pp.67-73, 情報処理学会, 2006.