

# PBLにおける学習習熟度を考慮した自主学習システムの提案

花田洋貴<sup>†1</sup> 大場みち子<sup>†2</sup>

**概要**：近年、実践的な ICT 人材の育成に PBL(Project Based Learning)を実施する高等教育機関が増えている。講義のみでは PBL に必要な前提知識を補いきれない。したがって、学生は PBL を実践しながら不足する知識をその都度自主学習をすることで補っている。自主学習では、学生の PBL 経験年数や専門知識の有無といった学習習熟度によって必要な情報が異なってくるため、自身の学習にふさわしい情報を適切に判断できない。このため、必要な情報を得るために時間がかかってしまい自主学習の効率がよくない。この問題を解決するために、PBL における自主学習の効率化を目指す。この目的を達成するために学生が閲覧した Web ページ中の情報から学生の習熟度を推定するアルゴリズムを提案する。提案したアルゴリズムを利用することで、習熟度に適した学習資料を推薦することができるかを確認するため、システム化と実験によって有用性を評価する。

**キーワード**：PBL, 自主学習, 習熟度, 推薦, 統計分析

## Proposal Self Learning Assistant System Considering Learning Level

HIROTAKA HANADA<sup>†1</sup> MICHIKO OBA<sup>†2</sup>

**Abstract**: Recently, higher education institutions bring up ICT human resources using PBL (Project Based Learning). Students can't compensate knowledge for PBL. For this reason, they do study themselves in order to compensate lack of knowledge in PBL. It is hard for them to determine accurate information because they are difference year of experience on PBL and knowledge of ICT. As a result, self learning is not efficient. Accordingly, our purpose is to be efficient of knowledge acquisition by self learning in PBL. In order to accomplish this purpose, we propose algorithm to estimate learning level using information of the web pages. To evaluate this algorithm, we develop a system and experiment the system.

**Keywords**: PBL, self learning, learning level, recommendation, Statistical analysis

### 1. はじめに

近年、情報系の高等教育機関において実践的な ICT 人材の育成にシステム開発型の PBL (Project Based Learning, 以下、システム開発型の PBL を単に PBL と記す) が成果をあげている[1]。PBL では、システム開発に関する基礎知識をもとに、課題を発見・解決するプロセスの中で実践的な知識・技術を習得することができる。学生は、PBL を実施する上で必要な前提知識を主に座学の講義で得る。しかし、座学の講義のみでは、PBL に必要な前提知識を全て補うことは難しい。講義で不足する知識を補うために、学生は PBL 実施前および実施中に必要なタイミングで自主学習を行う必要がある。自主学習では、講義スライド、Web ページ、関連書籍等 (以下、これらをまとめて学習コンテンツと記す) を用いて学習することで不足する知識を補う。しかし、自主学習において PBL の前提知識を効率よく学習できていないという現状がある。この原因として、学習コンテ

ツが膨大な情報量であるため、学生自身が、必要な情報に到達することに時間を要するからである。特に、PBL 経験年数やどれだけ専門知識を持っているかといった学生の学習習熟度によって適切な学習コンテンツが異なり、学生が自身の学習に適した資料をタイムリーに選択できないという問題がある。

本研究の目的を PBL における自主学習の効率化とする。目的を達成するために、学習習熟度の違いを考慮した自主学習システムを提案する。

### 2. 関連研究

自主学習を効率化するためには、学生にとってどのような学習システムが自主学習に適しているのかを知る必要がある。また、学生ごとに学習習熟度を判定する方法を検討しなければならない。これらに対して、以下の関連研究がある。

#### 2.1 学習コンテンツの推薦・提示による自主学習支援

藤原らは、PBL での自主学習支援を目的に、各学生の PBL の状況に合わせて、情報を推薦・提示するシステムを構築している[2]。具体的には、現在着手している開発工程

<sup>†1</sup> 公立はこだて未来大学 大学院  
Graduate School of Future University Hakodate  
<sup>†2</sup> 公立はこだて未来大学  
Future University Hakodate

やタスクに基づいて、必要あるいは関連のある学習コンテンツを推薦し提示する。図 1 に、学習コンテンツ推薦時の Web ページ例を示す。閲覧している Web ページ中の IT 用語に、各学生に推薦すべきコンテンツが存在した場合、その用語をハイライトし、マウスオーバーすると学習コンテンツへのリンクが表示される。各学生の PBL の状況に合わせて、適した情報を推薦・提示するというアプローチにより、それぞれが必要な情報にたどり着きやすくなり、自主学習の効率化に繋がる。しかし、学習習熟度に関しては考慮されていない。そのため、初めて PBL に参加する 1 年生に、前提知識が必要な Web ページが推薦・提示されることが起こりうる。本研究ではこの研究をもとに学習習熟度も考慮に入れた自主学習環境を構築する。

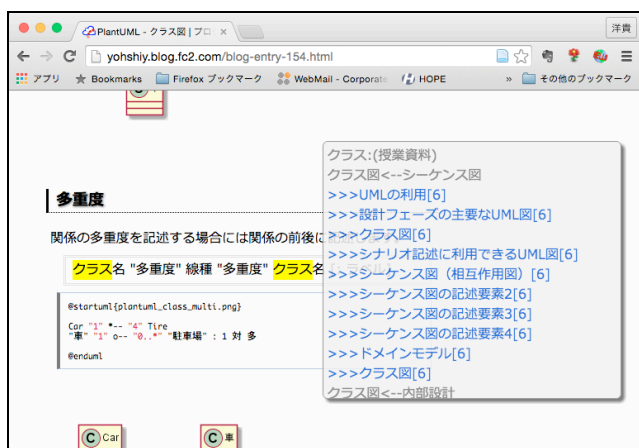


図 1 学習コンテンツの推薦

Figure 1 Recommend Learning Contents

## 2.2 学習習熟度判定手法

三好らは、閲覧している Web ページをもとにユーザの習熟度を判定する手法を提案している[3]。この手法では Web ページの難易度を推定し、易しいと推定されたページを閲覧しているユーザの習熟度は低く、難しいと推定されたページを閲覧しているユーザは習熟度が高いと判定される。また、閲覧している Web ページの順序を考慮することによって精度の向上を図っている。例えば、難しいページを見ていたユーザは習熟度が高いと判断される。しかし、そのページの次に易しいページを閲覧した場合、もとのページはユーザには難しすぎたため易しいページに遷移したと判断し、習熟度を当初の判定より低くするというアルゴリズムである。この研究では、Web ページの難易度推定に、“ある分野のアイテムを多く読んでいるユーザは、その分野について詳しい” (三好ら, 2012, p.1) という仮説を利用している。しかし、単にある分野のページへの閲覧数が多いからといって習熟度が高いとはいえない。特に、本研究で対象とする PBL に関する知識は、様々なページの概要を比較し、自分の習熟度にあったページを選んで読み始めるということが多々ある。例えば、習熟度の低い学生が「C 言

語」について学習することを考える。C 言語について書かれた Web ページはたくさんあるが、あるページを見た時、自分にとって理解しやすいページかどうかを判断し、理解しづらいと感じれば別のわかりやすいページを探しに行くといった具合である。したがって、この研究を本研究に当てはめると、正しく Web ページの難易度推定ができない。

Web ページ難易度推定については、劉らが、日本語を学習する外国人を対象とし、日本語テキスト中の文章をもとに、そのテキストが日本語能力検定の何級レベルなのかを推定する手法を提案している[4]。この手法を実現するために、文章中に日本語能力検定の何級レベルの語彙がどれだけ出現したかという割合を利用している。この割合を、あらかじめ求めておいた重回帰分析の回帰式に代入することで何級に相当するレベルのテキストかを求めている。この研究では、日本語の語彙のみに着目し割合を算出すれば良いが、PBL に必要な知識の中には、英語の語彙についても考慮する必要がある。そのため、この研究をそのまま本研究に適用することはできない。

## 3. 課題

目的を達成するためには、学習習熟度の判定を、PBL に適した手法で実施する必要がある。どういった PBL の知識を持っていけばどういった習熟度になるのかという基準をふまえた学習習熟度の推定アルゴリズムの構築を目指す。

## 4. 解決アプローチ

課題を解決するために、閲覧している Web ページの難易度をもとに、学生の習熟度を判定し、判定した習熟度にしたがって学生に必要な知識を、推薦、提示するシステムを構築する。Web ページの難易度推定には、情報処理技術者試験のシラバス[5]を利用する。シラバスにより分類された Web 文章中の用語から 3 つのレベルに区分する。

### 4.1 情報処理技術者試験のシラバスによる Web ページ難易度分類

システム開発に関して、自分の習熟度がどのくらいかを図るための試験として、情報処理技術者試験がある。図 2 は、情報処理技術者試験の試験区分の一覧である。情報処理技術者試験には、すべての社会人に共通に必要な基礎知識を問う IT パスポート試験(以下、IT パスポートと記す)、情報処理技術者として基本的な知識・技能を問う基本情報技術者試験(以下、基本情報と記す)、応用的な知識・技能を問う応用情報技術者試験(以下、応用情報と記す)、さらに高度な知識・技能を問う高度試験が 9 つに分かれて構成されている。図 2 の下段から上段に向かうにつれて試験の難易度は高まり、試験の難易度が高くなればなるほど、扱われる IT 用語も難しくなっていく。本研究では、文章中の用語を IT パスポートレベルの用語、基本情報レベルの用語、応用情報試験以上の用語という 3 つに振り分け、文中の用

語割合を算出することで Web ページの難易度推定に利用する。平成 28 年度から IT の安全な利活用の推進を目指した情報セキュリティマネジメント試験も開始されるが、本研究では対象外とする。

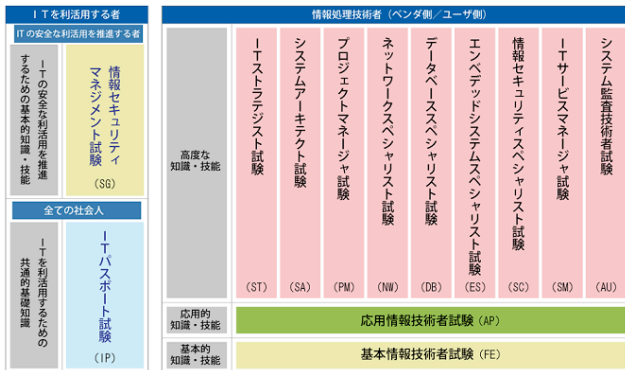


図 2 情報処理技術者試験の試験区分

Figure 2 Categories: Information Technology Engineers Examination

#### 4.2 Web ページ難易度推定の区分

劉らの研究では、求めたいデータである日本語能力検定何級相当かを推定するために、Web サイトに含まれる各レベルの語彙の割合を利用して、本研究では、「情報処理技術者試験のどのレベルに位置する学生が閲覧するのに適している難易度か」を求めたいデータとし、以下の3つのレベル区分に分類する。難易度は、レベル1が最も簡単で、レベルが上がるにつれて高い難易度のページになる。

- レベル1. 情報技術に関する基礎的な知識を持つ学生が閲覧するレベル
- レベル2. 基本的知識・技能をもち、実践的な活用能力を身につけた学生が閲覧するレベル
- レベル3. 応用的知識・技能をもち、高度 IT 人材としての方向性を確立した学生が閲覧するレベル

#### 4.3 Web ページ難易度推定手法

Web ページの難易度推定アルゴリズムのフローを図 3 に示し、以下で説明する。

- (1) Web ページ中の IT 用語の抽出  
文中に、IT パスポート、基本情報、応用情報以上の IT 用語がどのような割合で出現したかを算出する。
- (2) 抽出した IT 用語の分析  
算出された IT 用語の割合と、あらかじめ収集した標本データとを統計分析を用いて比較する。
- (3) 分析結果をもとに Web ページの難易度を推定  
統計分析の結果を利用して、Web ページの難易度として 4.2 節で定義した 3 つの区分のうちどのレベルに区分されるかを判断する  
劉らは、Web ページの難易度を推定する分析手法として、重回帰分析を利用した。重回帰分析は 1 つの目的変数と複

数の説明変数との関係を調べ、説明変数から目的変数を推定するための式(回帰式)を算出し利用する手法である。重回帰分析では、回帰式を求める際に利用した標本データに散らばりが多いと、適切な回帰式が求まらない。一方、判別分析[6]という分析手法がある。判別分析とは、対象者のデータからその個体がどの群に属するかを判定する手法である。この判別分析の中で、マハラノビス平方距離を利用した推定を行うと標本データの散らばりを考慮した判定が可能になる。したがって判別分析のほうが高い推定精度になるはずである。この仮説を検証するため、重回帰分析と判別分析どちらの推定精度が高いかを調べる予備実験を行う。予備実験をもとに、本研究で用いる分析手法を選定する。

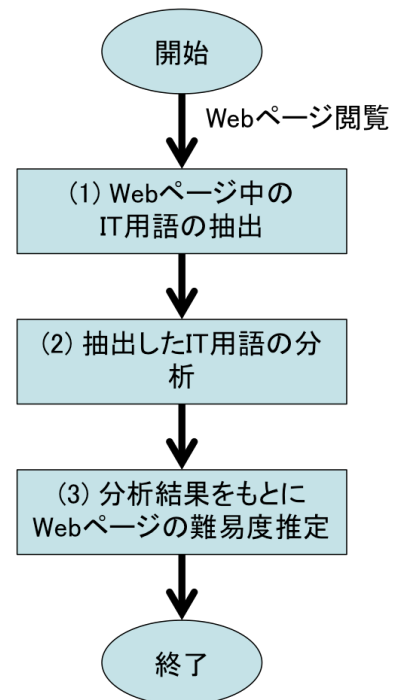


図 3 難易度推定アルゴリズムの概要  
 Figure 3 Summary of Estimate Algorithm

### 5. 予備実験

#### 5.1 目的

推薦精度の確認と本研究で用いる分析手法の選定を目的として行う。

#### 5.2 概要

各難易度 10 ページ分(計 30 ページ)の Web ページを標本データとし、標本データとは異なる 9 ページ分の実験データに対して Web ページの難易度推定を行う。あらかじめ筆者が判断した難易度と推定された難易度との一致率を重回帰分析と判別分析で比較する。重回帰分析での回帰式の算出には Microsoft Excel を利用し[7]、マハラノビス平方距離を利用した判別分析には R 言語[8]を利用した。また、Web ページ上の IT 用語抽出には MeCab[9]のユーザ辞書を利用

した。

### 5.3 標本データ

標本データとして、Web ページ中に出現した IT 用語の割合と、筆者が判断した Web ページの難易度を対応づけたものを利用する。ここでいう IT 用語とは、4.1 節で述べた IT パスポート、基本情報、応用情報以上の各試験で扱われる用語の割合のことである。Web ページの難易度は、4.2 節で述べた 3 つのレベルに分類している。標本データの例を以下の表 1 に示す。pageName は Web ページの URL、pageLevel は Web ページの難易度区分、IPWord、FEWord、APWord はそれぞれの IT 用語の割合(%)となっている。

表 1 標本データ例

Table 1 Example of sample data

pageName	pageLevel	IPWord	FEWord	APWord
<a href="http://itref.fc2">http://itref.fc2</a>	1	1.1601	5.8005	3.4803
<a href="http://itref.fc2">http://itref.fc2</a>	2	3.7439	6.4569	2.7672

### 5.4 実験対象データ

PBL で必要な知識の一つに UML[10]がある。UML について書かれている Web ページは、習熟度の低い人が学べるような入門レベルからある程度習熟している人を対象とした応用レベルまで多様である。PBL で必須の知識である点と、Web ページのレベルが多様で実験データを収集しやすいという点から UML に関する Web ページ 9 ページ分を実験対象データとして利用する。

### 5.5 予備実験結果

予備実験の結果を以下の表 2 に示す。重回帰分析と判別分析のどちらも 9 件中 4 件のページが正しく推定された。また重回帰分析は難易度レベル 2 に対しての推定精度が高く、判別分析では突出して高い推薦精度ではないが、重回帰分析では判定できなかったレベル 3 を判定することができている。

表 2 実験結果まとめ

Table 2 Summary of The Check Experiment

	重回帰分析	判別分析
難易度レベル 1	1/3	0/3
難易度レベル 2	3/3	2/3
難易度レベル 3	0/3	2/3

### 5.6 考察

仮説とは異なり、どちらの分析手法も同じ推定精度になった。これは、4.2 節で述べた Web ページの難易度区分を定性的に定義したことが問題だった。標本データ作成時に、各ページをどの難易度に区分すれば良いのか曖昧で、筆者の主観で区分を判断したことが原因だと考えられる。また、標本データの妥当性を確認するために、各標本データを、Web ページの難易度区分ごとに IT パスポート、基本情報、応用情報以上という 3 つを軸として 3 次元空間にプロット

した。プロットしたデータを図 4 に示す。図 4 から以下の 2 点が分かる。

- (1) 同じ難易度区分でもデータの散らばりが大きすぎるあまりにデータとしての散らばりが大きすぎるため、判別分析でも許容しきれなかったと考えられる。
- (2) 難易度区分は異なるが同じような位置に多くのデータが存在している

同じような位置に多くのデータが存在すると、実験データの IT 用語割合が少し違うだけで難易度区分が異なってしまうため、精度が高くならなかったと考えられる。

1,2 より、標本データ自体の妥当性も低いものになっていたことが分かった。したがってどの分析手法を本研究で用いるかについては、Web ページの難易度区分を定量的に判断できるものに定義しなおし、標本データの難易度区分を修正した後にもう一度同じ実験をすることで確認する必要がある。

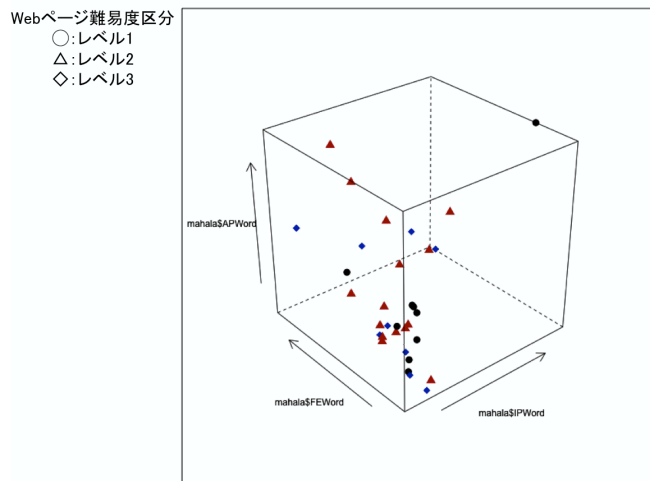


図 4 プロットした標本データ  
 Figure 4 Plots for Sample Data

## 6. 実験システムと評価

解決アプローチを組み込んだ実験システムを構築し、有用性を評価する。

### 6.1 実験システム概要

三好らの学習習熟度推定アルゴリズムを、3 章をもとに PBL に対応した形にし、学生の学習習熟度を推定する。推定された学習習熟度の学生ごとに適切な学習コンテンツを、藤原らのように推薦・提示するシステムである。システム構成図を以下の図 5 に示す。本稿 4 章で提案した習熟度推定アルゴリズムは、図中の 2 から 3 で利用する。実験システムのフローを以下に述べる。

- (1) 学生の閲覧ログ収集と解析  
 学生が閲覧している Web ページ文章中の、情報処理技術

者試験の各用語割合を取得し、分析を行う。

(2) 適切な学習コンテンツの推薦・提示

分析の結果から判断された学習習熟度に適した学習コンテンツを推薦・提示する。

(3) ページ閲覧順序を考慮した習熟度の再定義

別のページに移動した際、順序を考慮した習熟度の再定義を実施し、推薦・提示する学習コンテンツを変更する。

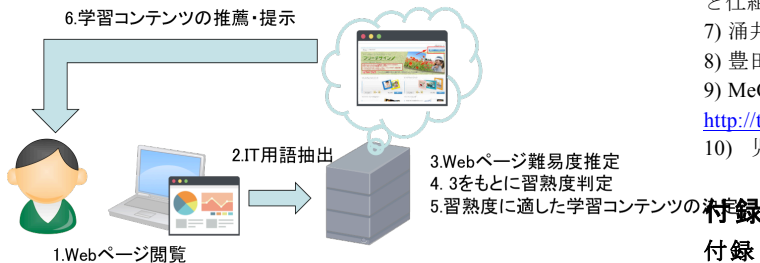


図 5 システム構成図

Figure 5 System Configuration Diagram

6.2 実験方法

構築したシステムを利用することで、学生の自主学習が効率化するかを検証するために実験を行う。PBL 経験年数が異なる学生 10 名前後を対象とし、システムを利用した場合と利用しなかった場合の推薦・提示された学習コンテンツの比較や、学生へのアンケートを実施することで評価を行う。

7. おわりに

PBL において、不足する知識は自主学習で補う必要がある。しかし、現状は、自主学習で PBL に必要な前提知識を効率よく学習できていない。この問題を解決するために、学生ごとの学習習熟度を考慮に入れた自主学習支援システムを提案した。学習習熟度を判定するために、Web ページの難易度推定を利用する。難易度推定には統計分析を利用する。本研究で利用する統計分析を決定するために予備実験を行った。しかし、予備実験では、定義した Web ページ難易度区分が定性的であることと標本データの信頼度が低いことが原因で適切な分析手法を決定することができなかった。今後は、Web ページ難易度区分を再定義した上で本研究にて利用する分析手法を選定する。その後、実験システムを構築し、PBL 経験年数の異なる学生 10 名前後に対して実験を行うことで有用性の検証を行っていく。

参考文献

1) 喜多一：「大学教育とプロジェクトマネジメント」, 国際プロジェクト・プログラムマネジメント学会 2006 年度春季研究発表大会 予稿集, pp.103-108, 2006.  
 2) 藤原哲, 大場みち子, 山口琢, 奥野拓, 伊藤恵：RDF とユーザプロフィールを用いた PBL 向け情報推薦手法, 日本ソフトウェア科学会第 31 回大会講演論文集, (2014).  
 3) 三好康夫, 濱田一伸, 鈴木一弘, 塩田研一, 岡本竜, 金西計英：

学習コンテンツ推薦に向けた 2 部ネットワーク分析に基づく習熟度と難易度の推定手法の提案, ARG W12 研究会第 1 回研究会, (2012).  
 4) 劉志宇, 内田理：日本語を学習する外国人を対象とした日本語テキスト難易度推定手法, 研究報告自然言語処理, vol. 2012-NL-205, no. 11, pp. 1-5(2012).  
 5) 情報処理技術者試験 試験区分一覧 [https://www.jitec.ipa.go.jp/1\\_1/seido/seido\\_gaiyo.html](https://www.jitec.ipa.go.jp/1_1/seido/seido_gaiyo.html)  
 6) 山口和範, 高橋淳一, 竹内光悦：よくわかる多変量解析の基本と仕組み, pp.74-78, 秀和システム, 2004.  
 7) 涌井良幸, 涌井貞美：実習多変量解析入門, 技術評論社, 2011.  
 8) 豊田秀樹：検定力分析入門, 東京図書, 2009.  
 9) MeCab: Yet Another Part-of-Speech and Morphological Analyzer <http://taku910.github.io/mecab/>  
 10) 児玉公信：UML モデリング入門, 日経 BP 社, 2008.

付録

付録として予備実験で用いた標本データと、テストデータを載せる。

付録 A.1 予備実験の標本データ

	pageName	pageLevel	IPWord	FEWord	APWord
1	<a href="http://itref.fc2">http://itref.fc2</a>	1	1.1601	5.8005	3.4803
2	<a href="http://itref.fc2">http://itref.fc2</a>	1	8.1818	3.6364	6.3636
3	<a href="http://itpro.nik">http://itpro.nik</a>	1	2.4474	2.7921	2.6543
4	<a href="http://itpro.nik">http://itpro.nik</a>	1	1.5092	1.7205	1.8714
5	<a href="http://itpro.nik">http://itpro.nik</a>	1	2.3184	2.5657	1.9474
6	<a href="http://itpro.nik">http://itpro.nik</a>	1	2.1039	2.5312	3.0572
7	<a href="http://itpro.nik">http://itpro.nik</a>	1	2.304	2.784	2.88
8	<a href="http://itpro.nik">http://itpro.nik</a>	1	1.9303	3.3898	2.2599
9	<a href="https://kotoba">https://kotoba</a>	1	2.1647	2.9519	0.8996
10	<a href="http://itref.fc2">http://itref.fc2</a>	1	3.7439	6.4569	2.7672
11	<a href="http://itref.fc2">http://itref.fc2</a>	2	3.3755	9.7046	4.6414
12	<a href="http://itref.fc2">http://itref.fc2</a>	2	5.7598	5.9436	3.9216
13	<a href="http://itref.fc2">http://itref.fc2</a>	2	5.1799	10.216	2.8777
14	<a href="http://itref.fc2">http://itref.fc2</a>	2	5.5448	7.2211	2.5145
15	<a href="http://itref.fc2">http://itref.fc2</a>	2	3.4894	11.662	5.326
16	<a href="http://itref.fc2">http://itref.fc2</a>	2	1.9153	7.0613	2.2828
17	<a href="https://www.o">https://www.o</a>	2	2.5109	1.8195	0.837
18	<a href="http://itpro.nik">http://itpro.nik</a>	2	1.9852	2.6316	2.4931
19	<a href="http://itpro.nik">http://itpro.nik</a>	2	1.8223	2.5513	2.4601
20	<a href="http://itpro.nik">http://itpro.nik</a>	2	1.8271	4.3672	1.6934
21	<a href="http://labo.ma">http://labo.ma</a>	3	2.5889	6.0407	1.0356
22	<a href="http://itref.fc2">http://itref.fc2</a>	3	0.8	3.8	4.8
23	<a href="http://home.a">http://home.a</a>	3	1.3128	1.3128	1.5982
24	<a href="http://www.ne">http://www.ne</a>	3	1.5725	3.4554	2.3795
25	<a href="http://akadem">http://akadem</a>	3	3.2439	2.8034	4.2451
26	<a href="http://applisio">http://applisio</a>	3	1.4616	1.8879	2.0706
27	<a href="http://e-words">http://e-words</a>	3	3.7406	5.4863	3.99
28	<a href="https://ja.wikip">https://ja.wikip</a>	3	1.0707	9.8501	3.8544
29	<a href="http://fukaoui.o">http://fukaoui.o</a>	3	1.0707	9.8501	3.8544
30	<a href="http://ameblo">http://ameblo</a>	3	1.7333	0.8	1.1333

付録 A.2 予備実験のテストデータ

	pageName	pageLevel	IPWord	FEWord	APWord
1	<a href="https://ja.wikipedia.org/">https://ja.wikip</a>	1	4.9083	6.4452	1.9584
2	<a href="http://objectc.com/">http://objectc</a>	1	3.49	5.8258	1.9511
3	<a href="http://e-words.jp/">http://e-words</a>	1	5.1325	9.106	3.4768
4	<a href="http://objectc.com/">http://objectc</a>	2	2.8898	6.7659	1.4189
5	<a href="http://itpro.nikkeibp.co.jp/">http://itpro.nik</a>	2	3.304	5.4623	2.1316
6	<a href="http://www.itn.co.jp/">http://www.itn</a>	2	3.2843	4.6569	1.9444
7	<a href="http://itpro.nikkeibp.co.jp/">http://itpro.nik</a>	3	2.7555	6.1424	2.5832
8	<a href="http://www.itn.co.jp/">http://www.itn</a>	3	1.9297	6.0088	1.5388
9	<a href="http://itpro.nikkeibp.co.jp/">http://itpro.nik</a>	3	1.9967	5.9425	1.9491