Paper:

Extraction of Web Site Evaluation Criteria and Automatic Evaluation

Peng Li* **and Seiji Yamada****

*Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology
J2, 4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan
E-mail:liorlee@nii.ac.jp
**National Institute of Informatics, SOKENDAI
2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-8430, Japan
[Received December 11, 2009; accepted March 8, 2010]

This paper proposes an automated web site evaluation using machine learning to extract evaluation criteria from existing evaluation data. Web site evaluation is a significant task because evaluated web sites provide information useful to users in estimating sites validation and popularity. Although many practical approaches have been taken to present possible measuring sticks for web sites, their evaluation criteria are manually determined. We developed a method to obtain evaluation criteria automatically and rank web sites with the learned classifier. Evaluation criteria are discriminant functions learned from a set of ranking information and evaluation features collected automatically by web robots. Experiments confirmed the effectiveness of our approach and its potential in high-quality web site evaluation.

Keywords: evaluation criteria extraction, web site evaluation

1. Introduction

Internet users number over 1.7 billion [1] and the number of web sites has grown exponentially. A November 2009 survey estimated 233 million [2] of these rich, highly updated information sources. Publishing on the Internet is easy, cheap, and fast. However, the web's very nature means that not all information is adequately or clearly attributed, or complete and trustworthy.

This had increased the use of web site evaluations in fields such as business, universities, government, and E-commerce. Based on Jupiter Research Reports, budget earmarked for web site evaluation products and services exceeded 1 billion US dollars annually by 2006.

No unified criteria currently exist that are fully trustworthy for web site evaluations. Each user has different criteria, and it is almost impossible to get a positive result using the same evaluation criteria on web sites of different categories. Moreover, the evaluation changes over time even for the same web site. Although many tools for evaluating the usability, accessibility, and security of web sites [3–6] have started to appear, the evaluation tasks are basically still done manually [7]. There are limitations as to how many web sites can be manually evaluated, and an automated evaluation system has been seen as a necessary step.

In this paper, we propose an automated method to extract evaluation criteria. This involves focusing on categorized web sites and setting evaluation criteria for individual categories applying machine learning to deal with ranking. We then use web robots to collect the features of the evaluation objectives and evaluate them automatically based on their features and the criterion of the corresponding category.

Establishing evaluation criteria requires that evaluation data be obtained regularly and automatically from the general public – a task conventionally done by questionnaires. This has changed however, with the advent of registration ranking sites – real-time web site ranking services based on user recommendations and votes. The most popular of these, such as *SiteRank*[1], have hundreds of thousands of web sites registered and users make categories, register new sites, vote, and write comments frequently. In this research, we use categorized ranking information from registration ranking sites to set up web sites evaluation criteria. In other words, we treat automated web site evaluation as a classification learning problem in which discriminant functions are learned by categorized ranking information used as training data. We compared three major ranking algorithm – Ranking SVM [8], Prank [9], and Multiclass SVM [10] – to find the one best adapted to solving the ranking problem.

2. Background

Web site evaluation task requires appropriate evaluation measures. Researches such as Web WISDOM [11] and WebMac [12] use indexes from library and information science findings as meta-information to create evaluation measures, although these do not particularly target automation.

Aside from the approach of the questionnaire sur-

1. http://siterank.org/jp/

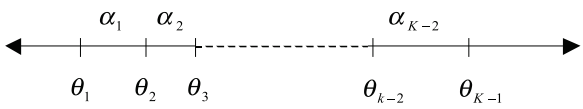

Fig. 1. Ranking.

vey, Velayathan and Yamada [13] proposed automatically learning evaluation criteria based on user browsing behavior, but evaluation objectives were web pages, not web sites, and users had to browse evaluated objectives to ascertain browsing behavior.

Tools focusing on evaluating web site usability, accessibility, and security are roughly divided into (1) server performance analysis, (2) usage analysis based on log data, (3) guideline compliance checking, (4) navigation text analysis, (5) cyber agent navigation simulation [14]. Automatically evaluating web sites by these tools, however, involves evaluation criteria predetermined by experts.

3. Evaluation Criteria Extraction

We define evaluation criteria as discriminant functions learned from a set of ranking information. In this section, we introduce algorithms that solve the ranking problem and detail evaluation features and procedures.

3.1. Ranking Algorithms

Our ranking problem uses ranking information as both input and output. Algorithms such as Ranking SVM, Prank, and Multiclass SVM used to solve this kind of problem are detailed below.

3.1.1. Ranking SVM

As presented graphically in **Fig. 1**, the distance from hyperplane w of datapoint x is mapped onto one-dimensional (1D) space in which $\theta_1,\ldots,\theta_{K-1}$ are different thresholds against which distance is compared. Data point x having rank i satisfies $\theta_{i-1} < w^T x_j < \theta_i$. Checking which section $w^T x_j$ goes into helps determine the rank of x. Define $\alpha_i = \theta_{i+1} - \theta_i$; $1 \le i \le K-1$. No lower bound exists for data items belonging to rank 1 and all data items belonging to rank K have no upper bound. Note that data point x_j having rank $i > 1$ satisfies the following:

$$w^T x_j > \theta_{i-1}; w^T x_j + \alpha_{i-1} > \theta_i; w^T x_j + \sum_{k=i-1}^{K-2} \alpha_k > \theta_{K-1}. \quad (1)$$

Similarly, for an example with rank $i < K$, we have the following:

$$w^T x_j < \theta_i; w^T x_j + \alpha_i < \theta_{i+1}; w^T x_j + \sum_{k=i}^{K-2} \alpha_k < \theta_{K-1}. \quad (2)$$

Based on these observations, define $\bar{w} = [w, \alpha_1, \alpha_2, \ldots, \alpha_{K-2}]$ and for example x_j with rank $1 < i < K$, define $\bar{x}_j^+$ and $\bar{x}_j^-$ as $n+K-2$ dimensional vectors as follows:

$$\bar{x}_j^+[l] = \begin{cases} x_j[l] & 1 \le l \le n \\ 0 & n < l < n+i-1 \\ 1 & n+i-1 \le l \le n+K-2 \end{cases}$$

;

$$\bar{x}_j^-[l] = \begin{cases} x_j[l] & 1 \le l \le n \\ 0 & n < l < n+i \\ 1 & n+i \le l \le n+K-2 \end{cases} \quad (3)$$

This leads to the following:

$$\bar{w}^T \bar{x}_j^+ > 0 \Rightarrow w^T x_j > \theta_{i-1}; \bar{w}^T \bar{x}_j^- < 0 \Rightarrow w^T x_j < \theta_i. \quad (4)$$

Once $\bar{x}_j^+$ and $\bar{x}_j^-$ are defined, the ranking problem is simply reduced to learning classifier $\bar{w}$ in $n+K-2$ dimensional space so that $\bar{w}^T \bar{x}_j^+ > 0$ and $\bar{w}^T \bar{x}_j^- < 0$. This standard classification problem has at most 2m training examples, half of which have label $+1$ (examples $\bar{x}_j^+$) and the other half label -1 (examples $\bar{x}_j^-$).

Many real-world problems may not yield a linear function powerful enough to learn data ranking, so nonlinear mapping $\phi()$ is used over the original feature vector. If x_j has rank i, then Eq. (4) is satisfied. However:

$$\kappa(\bar{w}, \bar{x}_j^+) = \phi(\bar{w})^T \phi(\bar{x}_j^+) > 0 \Rightarrow \phi(w)^T \phi(x_j) > \theta_{i-1} \quad (5)$$

is not satisfied due to the nonlinearity of $\phi()$. To get around this problem, a new kernel function has to be defined. For kernel function K and corresponding mapping ϕ, define new kernel function $\bar{\kappa}$ with the corresponding mapping $\phi()$ as follows:

$$\phi(\bar{w}) = [\phi(w), \bar{w}[n+1:n+K-2]]$$

;

$$\phi(\bar{x}) = [\phi(x), \bar{x}[n+1:n+K-2]]. \quad (6)$$

A kernel trick becomes available with the new kernel function:

$$\bar{\kappa}(\bar{w}, \bar{x}) = \kappa(w, x). \quad + \bar{w}\ [n+1:n+K-2] \quad \bar{x}\ [n+1:n+K-2] \quad (7)$$

3.1.2. Prank

Prank, a perceptron-based online learning algorithm, trains a perceptron model, retaining both weight vector w and threshold vector b, with the objective of finding a perceptron model that projects all instances into k subintervals defined by b. **Fig. 2** shows the Prank algorithm.

3.1.3. Multiclass SVM

Multiclass SVM, a generalization of two-class SVM systems, deals with multiclass problems mainly by breaking down single multiclass problems into multiple binary problems, each yielding a binary classifier assumed to produce an output function giving relatively large values

Initialize: Set $w^1 = 0, b_1^1, \ldots, b_{k-1}^1 = 0, b_k^1 = \infty$
Loop: For $t = 1, 2, \ldots, T$
- Get a new rank-value $x^t \in \mathbb{R}^n$.
- Predict $\hat{y}^t = \min_{r \in \{1,\ldots,k\}} \{r : w^t \cdot x^t - b_r^t < 0\}$.
- Get a new label y^t.
- If $\hat{y}^t \neq y^t$ update w^t (otherwise set $w^{t+1} = w^t, \forall r : b_r^{t+1} = b_r^t$):
  1. For $r = 1, \ldots, k-1$: If $y^t \leq r$ Then $y_r^t = -1$
     Else $y_r^t = 1$.
  2. For $r = 1, \ldots, k-1$: If $(w^t \cdot x^t - b_r^t) y_r^t \leq 0$ Then $\tau_r^t = y_r^t$
     Else $\tau_r^t = 0$
  3. Update $w^{t+1} \leftarrow w^t + (\sum_r \tau_r^t) x^t$.
     For $r = 1, \ldots, k-1$ update: $b_r^{t+1} \leftarrow b_r^t - \tau_r^t$

Output: $H(x) = \min_{r \in \{1,\ldots,k\}} \{r : w^{T+1} \cdot x - b_r^{T+1} < 0\}$.

Fig. 2. Prank algorithm.

to examples from the positive class and relatively small values to examples belonging to the negative class. To deal with k-class classification, for example, it generates set of 2-class classification SVMs $f^1, \ldots, f^k$ that discriminate between one class and the rest, then classifies multi-classes based on Eq. (8) to integrate $f^1, \ldots, f^k$.

$$\arg\max_{j=1,\ldots k} g^j(x). \quad \ldots \quad (8)$$

$g^j(x)$ is expressed as follows:

$$g^j(x) = \sum_{i=1}^{l} y_i \alpha_i^j \cdot k(x, x_j) + b^j. \quad \ldots \quad (9)$$

The final complex function is as follows:

$$f^j(x) = sgn(g^j(x)). \quad \ldots \quad (10)$$

3.2. Evaluation Features

We use ranking information from registration ranking sites as training data. Each data consists of multiple evaluation features and their values, and the data value is the rank determined by ranking sites. Using discriminant functions learned from training data, we rank unranked web sites based on their features. Data from these unranked web sites is test data.

Sinha et al. [15] and Ivory et al. [16] studied "best" web sites selected by experts and users to identify common features. These include content, structure and navigation, visual design, functionality, and interactivity. Palmer [17] concluded from three consecutive tests that web site success depends on speed, navigation, content, interactivity, and response. These researches deal with both subjective and objective user features. Our goal, however, is to automate evaluation, so the features we need must be objective and obtained automatically from web servers. **Table 1** lists 31 features categorized in 10 fields currently used in our research and detailed below.

(i) Global Link Popularity shows a web page's popularity by examining the link structure of the web. We use Google PageRank here.

(ii) Freshness is the contrast of new and old information. We use update frequency and the amount of text information changed as a guide.

Table 1. Ten fields of evaluation features.

Field	The number of features
Top page's global link popularity	1
Freshness	2
Indexable text information	1
Multi-media content	4
Accuracy of spelling and grammar	1
Accuracy of HTML documents	4
Content security	5
Content constitution	4
Design	5
Others	4

(iii) Indexable Text Information is the amount of visible HTML text information obtained by calculating the size of HTML source files with tag information removed.

(iv) Multimedia Content means the number of files including images, videos, audios and flashes.

(v) Spelling and Grammar Accuracy counts the number of misspellings.

(vi) HTML Document Accuracy checks features such as whether character code and image size are specified.

(vii) Content Security checks features related to the control of cache, scripts, and web robots.

(viii) Content Constitution uses the proportion of multimedia content to text information as a guide.

(ix) Design considers information on BGM, background image, frames, style sheets, etc.

(x) Others includes the number of links on the top page, the number of links overall, whether the page author is specified, and whether the page description is specified.

These 31 features were selected under the restriction of currently available web server technology. Although other features will become available as web technology advances, we believe these features are appropriate for our purpose.

3.3. Automated Web Site Evaluation

As shown in **Fig. 3** overview of our web site evaluation system, in the learning phase, categorized ranking information from registration ranking sites and evaluation features collected by web robots are input to the ranking learning machine to learn categorized discriminant functions. In the test phase, ranks of unranked web sites are predicted automatically based on site's features and corresponding discriminant functions.

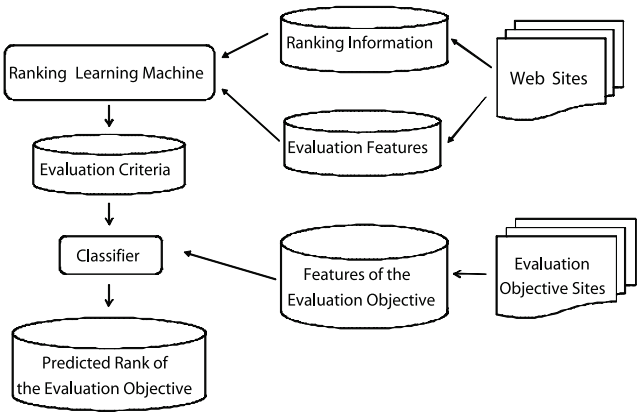

Fig. 3. System overview.

4. Experiments

We compared ranking algorithms Ranking SVM, Prank, and Multiclass SVM experimentally to find the one best adapted to this research. We then verified the adequacy of extracted evaluation criteria and studied the correlation between evaluation features and classification results and compared our evaluation to a web site importance estimation approach.

The dataset consisted of 735 web sites in seven categories in which each category had web sites divided into five equal groups and had labels assigned ranging from 1, Excellent, to 5, Not expected. The 31 evaluation features discussed in Section 3.2 were selected. LibSVM2.84 was used as the SVM tool.

4.1. Evaluation Measures

We used Mean Absolute Error (MAE) – a statistical quantity used to measure how close forecasts or predictions are to eventual outcomes – to evaluate performance. MAE is given by:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N} |p_i - q_i|}{N}. \qquad (11)$$

q_i is prediction, p_i the true value, and N the number of items – the smaller the MAE, the higher the classification accuracy.

4.2. Data Preprocessing and Kernel Selection

Values of the evaluation features vary from 0 to millions, so we scaled data pre-experimentally to [0,1]. LibSVM provides four kernels – linear, polynomial, radial basis function, and sigmoid. Experimentally studying kernel accuracy, we decided to use the RBF kernel due to its superior accuracy and reprogrammed part of the kernel source code to apply the kernel trick for Ranking SVM.

We did a grid search using cross-validation to find the best value of RBF kernel parameters – c and γ – for each of the seven web site categories.

Table 2. Experimental algorithm comparison.

	MAE	Minimum number of adjacent transpositions
Ranking SVM	0.78	132.7
Multi-class SVM	0.93	217.4
Prank	1.12	264.5

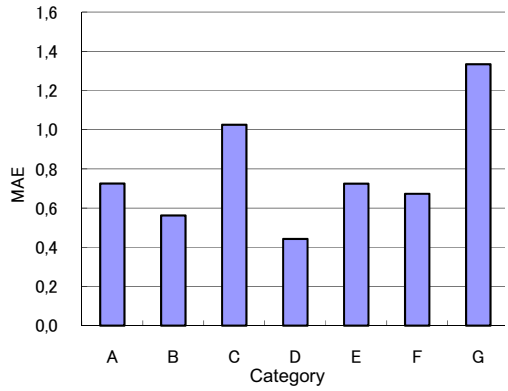

Fig. 4. Evaluation experiment results. (The y-axis represents the values of MAE, and the x-axis corresponds to categories A: computer science, B: music, C: health and diet, D: movies, E: sports, F: games, and G: fashion.)

4.3. Algorithm Comparison

We experimentally compared Ranking SVM, Prank, and Multiclass SVM, getting the results shown in **Table 2**. To get detailed information in addition to MAE, we evaluated the three algorithms by calculating the minimum number of adjacent transpositions [18] needed to bring the predicted value to ground truth. As **Table 2** shows, Ranking SVM significantly outperformed Prank and Multiclass SVM, implying that the ranking problem in this research is nonlinear.

4.4. Criteria Extraction Accuracy Evaluation

In an experiment to verify the adequacy of extracted evaluation criteria, we conducted 10-fold cross-validation in each category, as shown in **Fig. 4**, to minimize the influence of biased data. Average MAE is 0.78 and standard deviation is 0.3, meaning that, in most cases, the predicted value is either equal to the true value or shifted one rank from it. We looked at studies [19–21] that also use MAE with a 5-level evaluation and found that keeping MAE under 1 is sufficiently effective.

4.5. Feature Contribution Ratio Experiment

It is important for users and web site authors to know the correlation between evaluation features and classification results. Since the ranking problem in this research is nonlinear and we use a kernel trick in Ranking SVM, we cannot obtain the feature contribution ratio directly from the learned discriminant functions, so we used a wrapper [22], widely used in feature selection, to determine the feature contribution ratio. **Table 3** shows the top five

Table 3. Top five features by category.

	computer science
1	Indexable text information
2	Percentage of updated text information
3	Number of links (site)
4	Specification of author information
5	Top page's PageRank
	music
1	Number of links (top page)
2	Number of audio files
3	Number of links (site)
4	Top page's PageRank
5	Percentage of audio files
	health and diet
1	Indexable text information
2	Number of image files
3	PR for search robots
4	Specification of image size
5	Number of links (site)
	movies
1	Number of links (site)
2	Number of image files
3	Number of links (top page)
4	Percentage of image files
5	Specification of author information
	sports
1	Number of image files
2	Number of links (top page)
3	Indexable text information
4	Specification of background image
5	Top page's PageRank
	games
1	Number of links (site)
2	Number of image files
3	Specification of background image
4	Number of links (top page)
5	Indexable text information
	fashion
1	Indexable text information
2	Percentage of updated text information
3	Number of image files
4	Specification of cache expiration date
5	Specification of font code

most contributed evaluation features in each category. We concluded that feature contribution ratios vary with the category.

4.6. Comparison to Existing Importanc Estimation

To verify its effectiveness, we compared our evaluation to an existing web site importance estimation approach. We chose PageRank, the most common importance estimation, for our comparison. Classification accuracy was compared for each category, as shown in **Table 4**. Our evaluation shows better classification accuracy than PageRank in all categories.

Table 4. MAE comparison to PageRank.

Category	Proposal	PageRank
Computer science	0.73	1.57
Music	0.56	2.12
Health and diet	1.03	1.86
Movies	0.44	1.73
Sports	0.72	2.23
Games	0.67	2.57
Fashion	1.33	1.48
Average MAE	0.78	1.94

5. Discussion

5.1. Category Influence

As shown in **Fig. 4**, differences in evaluation accuracy between categories are great. Categories with low accuracy exist for two reasons:

(1) It is difficult to set unified evaluation criteria for some categories due to their characteristics, e.g., users visiting health and diet web sites may have very different objectives – looking for health advice, food, recipes, or even body-building equipments – from users visiting movie and music category web sites, meaning that different purposes lead to different evaluation criteria.

(2) Important features that are the key to determining evaluation criteria are not acquired experimentally, e.g., if beautiful high-resolution images are preferred in the fashion category, it is difficult to set a high accuracy criterion without a feature describing it. More evaluation features are likely to become available with continuing improvement of web robots and web technology. With more adapted features, things improve for the low-accuracy categories.

5.2. Training Data Quantity and Quality

Our data came from over 700 web sites. To get enough for each category, we collected data from multiple registration ranking sites, although this may potentially decrease evaluation accuracy due to changes in user group and the web site quality bias. With further development of registration ranking sites, enough data will become available from a specific site, at which point evaluation accuracy is expected to improve. With registration ranking site categories becoming more specific, we expect evaluation accuracy to increase as categories become more specialized.

5.3. Evaluation Features and Contribution Ratio

Table 3 shows that features such as indexable text information and numbers of links play an important role in most categories. However, the fact that some features are effective only for corresponding categories shows that features reflect category characteristics. Features in **Table 3**, for example, contribute much to criteria extraction accuracy, although some features hardly have any effect or even an opposite effect – in which case it may be advantageous to remove such features at this time. Because

evaluation criteria change over time, however, and some features may eventually become effective, we do not remove any existing features.

6. Conclusions

Improving the quality of web content is imperative for further web development, because good measuring sticks are needed for web sites. Many practical approaches have been made in this regard, but have required manual intervention, preventing dealing with the entire web.

The automated web site evaluation we have proposed involves defining web site evaluation as web site ranking and applying machine learning to extract evaluation criteria. We compared three ranking algorithms, finding Ranking SVM to be the best adapted to solving the ranking problem. Our evaluation experiments confirmed the feasibility of extracted evaluation criteria. We also compared our evaluation to an existing web site importance estimation approach and verified that our evaluation has the advantage of evaluation accuracy. We also conducted feature contribution experiments to determine the correlation between evaluation features and classification results.

We have presented a practical approach to automatically evaluating web sites that has high potential in high-quality web site evaluation.

References:

[1] http://www.internetworldstats.com/

[2] http://news.netcraft.com/

[3] S. Ssemugabi and R. de Villiers, "A comparative study of two usability evaluation methods using a web-based e-learning application," Proc. of 2007 Annual Research Conf. of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries, pp. 132-142, 2007.

[4] A. Aizpurua, M. Arrue, M. Vigo, and J. Abascal, "Transition of accessibility evaluation tools to new standards," Proc. 2009 Int. Cross-Disciplinary Conf. on Web Accessibility, pp. 36-44, 2009.

[5] E. Velleman, C. Strobbe, J. Koch, C. A. Velasco, and M. A. Snaprud, "Unified Web Evaluation Methodology Using WCAG," Proc. of the 4th Int. Conf. on Universal Access in Human-Computer Interaction, Vol.4556, pp. 177-184, 2007.

[6] L. Falk, A. Prakash, and K. Borders, "Analyzing Websites for User-Visible Security Design Flaws," Proc. of the 4th Symposium on Usable Privacy and Security, pp. 117-126, 2008.

[7] O. Chapelle and Y. Zhang, "A dynamic bayesian network click model for web search rank," Proc. of the 18th Int. Conf. on World Wide Web, pp. 1-10, 2009.

[8] S. Rajaram, A. Garg, X. S. Zhou, and T. S. Huang, "Classification approach towards ranking and sorting problems," Proc. of the 14th European Conf. on Machine Learning, pp. 301-312, 2003.

[9] K. Crammer and Y. Singer, "Pranking with ranking," Advances in Neural Information Processing Systems 14, MIT Press, Vol.1, pp. 641-647, 2002.

[10] E. L. Allwein, R. E. Schapire, and Y. Singer, "Reducing Multiclass to Binary: A Unifying Approach for Margin Classifiers," J. of Machine Learning Research, Vol.1, pp. 113-141, 2000.

[11] J. E. Alexander and M. A. Tate, "Web WISDOM: How to evaluate and create information quality on the Web," Lawrence Erlbaum, Hillsdale, 1999.

[12] M. P. Arnone and R. V. Small, "WWW motivation mining finding treasures for teaching evaluation skills grade 1-6," Linworth Publishing, Worthington, 1999.

[13] G. Velayathan and S. Yamada, "Behavior-based Web page evaluation," J. of Web Engineering, pp. 222-243, 2007.

[14] M. Y. Ivory, "An Empirical Approach to Automated Web Site Evaluation," J. of Digital Information Management, Vol.1, No.2, pp.75-102, 2003.

[15] R. Sinha, M. Hearst, and M. Ivory, "Content or Graphics?: An Empirical Analysis of Criteria for Award-Winning Websites," Proc. of the 7th Conf. on Human Factors and the Web, 2001.

[16] M. Y. Ivory and M. A. Hearst, "Statistical Profiles of Highly-Rated Web Sites," Proc. of the 20th ACM Conf. on Human Factors in Computing Systems, 2002.

[17] W. P. Palmer, "Web Site Usability, Design, and Performance Metrics," Information Systems Research, Vol.13, No.2, 2002.

[18] G. Lebanon and J. Lafferty, "Cranking: Combining rankings using conditional probability models on permutations," Proc. of the 19th Int. Conf. on Machine Learning, pp. 363-370, 2002.

[19] H. Takashima, H. Yamagishi, and S. Hirasawa, "An Improved Method of Collaborative Filtering with Predicting Unobserved Values," FIT Japan, A-008, 2005.

[20] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," Proc. of the 10th Int. Conf. on World Wide Web, pp. 285-295, 2001.

[21] P. Li and S. Yamada, "A Movie Recommender System Based on Inductive Learning," Proc. of IEEE Conf. on Cybernetics and Intelligent Systems, Vol.1, pp. 318-323, 2004.

[22] R. Kohavi, and G. H. John, "Wrappers for Feature Subset Selection," Artificial Intelligence, Vol.97, No.1-2, pp. 273-324, 1997.

Name:
Peng Li

Affiliation:
Tokyo Institute of Technology

Address:
4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan

Brief Biographical History:

2002- Master's Student, Tokyo Institute of Technology
2004- Doctor's Student, Tokyo Institute of Technology
2010- Researcher, Productive Leader Incubation Platform, Tokyo Institute of Technology

Main Works:

• "Web Site Evaluation Criteria Extraction and Evaluation Automation," IEEJ Japan, Vol.130, No.1, 66-74, Jan. 2010.

Membership in Academic Societies:

• The Japanese Society for Artificial Intelligence (JSAI)

Name:
Seiji Yamada

Affiliation:
National Institute of Informatics, the graduate university for advanced studies (SOKENDAI)

Address:
2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-8430, Japan

Brief Biographical History:

1989- Research Associate, Osaka University
1991- Lecturer, Osaka University
1996- Associate Professor, Tokyo Institute of Technology
2002- Professor, National Institute of Informatics, SOKENDAI

Main Works:

• human-agent interaction, intelligent interactive systems, interactive machine learning

Membership in Academic Societies:

• The Institute of Electrical and Electronics Engineers, Inc. (IEEE)
• Association for the Advancement of Artificial Intelligence (AAAI)
• The Japanese Society for Artificial Intelligence (JSAI)