Chapter 5 Research Experiment

5.1 Experimental Design

A set of experiments designed to verify the workload model is presented in this study by using a prototype system. The prototype system can connect to a web search service to do a proper search. Following our survey, we chose Yahoo! Web Search APIs because they provide the largest number of APIs and the most performance metrics.

5.1.1 Experimental Specification

Our set of experiments consists of ten basic and synthetic tests for benchmarking. Each test includes a case and a specific function of the prototype system. After finishing the experiments, the workload model is completely verified. We specify the corresponding selection of the prototype system and follow specifications in execution of each test. The specifications of each test consist of a page model selection, a query model selection, a keyword and the corresponding parameters for web search service APIs. The details of each test are described in the following sections.

(1) Test1:Single Page- URL

The first test suite is Single Page-URL. The purpose is to test "Anchor Text" of Single Page. To complete the test, the corresponding selection in the page model and the query model is chosen. The parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. "URL" in the Page Model is selected. In addition, this test is a kind of "topic distillation" query, so we also select "Topic Distillation" in the Query Model. The keyword is "SOA". The two parameters of the Yahoo! Search APIs are set to "query=inurl:SOA" and "Results=100". Table 5.1 shows the specifications of the test suite.

1	<u> </u>
Page Model: Single Page	URL
Query Model: Query Type	Topic distillation
Keyword	SOA
Yahoo! Search API parameter	query = inurl: SOA
Yahoo! Search API parameter	Results =100

Table 5.1: The Test Specifications of Single Page-URL

(2) Test2:Single Page-Font Size, Font Color, Frame, Meta, Table

The second test suite is Single Page-Font Size, Font Color, Frame, Meta, and Table. The purpose is to test "Font Size", "Font Color", "Frame", "Meta" and "Table" of Single Page. To complete this test, the corresponding selection in the page model and the query model should be chosen. The parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. The options for "Font Size", "Font Color", "Frame", "Meta" and "Table" are selected in the Page Model. This test is a kind of "topic distillation" query so we also select "Topic Distillation" in the Query Model. The keyword is "SOA". The two parameters of the Yahoo! Search APIs are set to "query = SOA" and "Results = 100". Table 5.2 shows the specifications of the test suite.

Page Model: Single Page	Font Size, Font Color, Frame, Meta, Table
Query Model: Query Type	Topic distillation
Keyword	SOA
Yahoo! Search API parameter	query = SOA
Yahoo! Search API parameter	Results = 100

Table 5.2: The Test Specifications of Single Page-Font Size, Font Color, Frame, Meta and Table

(3) Test3: Single Page-Title

The third test suite is Single Page-Title. The purpose is to test "Title" of Single Page. To complete this test, the corresponding selection in the page model and the query model should be chosen. And the parameters that are required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. "Title" option is selected in the Page Model. In addition, this test is a kind of "Topic distillation" query so we also select the "Topic Distillation" in the Query Model. The keyword is "SOA". The two parameters of the Yahoo! Search APIs are set to "query = i ntitle: SOA" and "Results = 100". Table 5.3 shows the specifications of the test suite.

Page Model: Single Page	Title
Query Model: Query Type	Topic distillation
Keyword	SOA
Yahoo! Search API parameter	query = intitle: SOA
Yahoo! Search API parameter	Results = 100

Table 5.3: The Test Specifications of Single Page-Title

(4) Test4:Multi Page-Company

The fourth test suite is Multi Page-Company. The purpose is to "Company" of Multi Page. To complete this test, the corresponding selection in the page model and the query model should be chosen. And the parameters that required by the Yahoo! Web Search API and keywords should be

entered to generate the web search script. The "Company" option is selected in the Page Model. In addition, this test is a kind of "Topic distillation" query so we also select the "Topic Distillation" in the Query Model. The keyword is "SOA". The three parameters of the Yahoo! Search APIs are set to "query = intitle: SOA", "site = www.ibm.com" and "Results = 100". Table 5.4 shows the specifications of the test suite.

Page Model: Multi-Page	Company
Query Model: Query Type	Topic distillation
Keyword	SOA
Yahoo! Search API parameter	query=SOA
Yahoo! Search API parameter	site=www.ibm.com
Yahoo! Search API parameter	Results=100

Table 5.4: The Test Specifications of Multi Page-Company

(5) Test5:Query Type-Homepage Finding

The fifth test suite is Query Type-Homepage Finding. The purpose is to test "Homepage Finding" of Query Type to find the homepage of a website. To complete this test, the corresponding selection in the query model should be chosen. And the parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. The "Homepage Finding" is selected in the Query Model. The keyword is "ibm". The two parameters of the Yahoo! Search APIs are set to "query = hostname:www.IBM.com" and "Results = 100". Table 5.5 shows the specifications of the test suite.

1	
Query Model: Query Type	Homepage finding
Keyword	ibm
Yahoo! Search API parameter	query = hostname:www.IBM.com
Yahoo! Search API parameter	Results = 100

Table 5.5: The Test Specifications of Query Type- Homepage Finding

(6) Test6: Query Type- Named page Finding

The sixth test suite is Query Type- Named page Finding. The purpose is to test "Named page Finding" in the Page Model to find the specific page. To complete this test, the corresponding selection in the query model should be chosen. And the parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. The "Named page Finding" option is selected in the Query Model. The keyword is the URL "help.yahoo.com/help/us/ysearch/tips/tips-04.html". The two parameters of the Yahoo! Search APIs are set to "query = url: help.yahoo.com/help/us/ysearch/tips/tips-04.html" and "Results = 100". Table 5.6 shows the specifications of the test suite.

Query Model: Query Type	Named page finding
Keyword	help.yahoo.com/help/us/ysearch/tips/tips-04.html
Yahoo! Search API parameter	query=url:
	help.yahoo.com/help/us/ysearch/tips/tips-04.html
Yahoo! Search API parameter	Results = 100

Table 5.6: The Test Specifications of Query Type- Named page Finding

(7) Test7: Query Type-Topic Distillation

The seventh test suite is Query Type-Topic Distillation. The purpose is to test "Topic Distillation" of Query Type to find the relative pages of a keyword. To complete this test, the corresponding selection in the query model should be chosen. And the parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. The "Topic Distillation" option is selected in the query model. The keyword is "SOA". The two parameters of the Yahoo! Search APIs are set to "query = SOA" and "Results = 100". Table 5.7 shows the specifications of the test suite.

Tuble 5.7. The fest specifications of Query Type Tople Distinution	
Query Model: Query Type	Topic distillation
Keyword	SOA
Yahoo! Search API parameter	query = SOA
Yahoo! Search API parameter	Results = 100

Table 5.7: The Test Specifications of Query Type-Topic Distillation

(8) Test8:Link Structure-Authority-Hub

The eighth test suite is Link Structure-Authority-Hub. The purpose is to test "Authority-Hub" of Link Structure to find the relative pages that link to the authoritative page. To complete this test, the corresponding selection in the query model should be chosen. And the parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. the "Authority-Hub" option is selected in the query model. In addition, this test is a kind of "Topic distillation" query so we also select the "Topic Distillation" option in the query model. The keyword is "SOA". The two parameters of the Yahoo! Search APIs are set to "query = link:ibm.com.tw+soa" and "Results = 100". Table 5.8 shows the specifications of the test suite.

Table 5.8: The Test Specifications of Link Structure-Authority-Hub

Query Model: Link Structure	Authority-Hub
Query Model: Query Type	Topic distillation

Keyword	SOA
Yahoo! Search API parameter	query = link:ibm.com.tw+soa
Yahoo! Search API parameter	Results = 100

(9) Test9:Similarity-TLS

The ninth test suite is Similarity- TLS. The purpose is to test "Similarity- TLS". To complete this test, the corresponding selection in the query model should be chosen. And the parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. The "TLS" option is selected in the query model. In addition, this test is a kind of "Topic distillation" query so we also select the "Topic Distillation" option in the query model. The keyword is "Service Oriented Architecture". The three parameters of the Yahoo! Search APIs are set to "query= Service Oriented Architecture", "type=phrase" and "Results=100". Table 5.9 shows the specifications of the test suite.

1 5	
Query Model: Similarity	TLS
Query Model: Query Type	Topic distillation
Keyword	Service Oriented Architecture
Yahoo! Search API parameter	query = Service Oriented Architecture
Yahoo! Search API parameter	type = phrase
Yahoo! Search API parameter	Results = 100

Table 5.9: The Test Specifications of Similarity- TLS

(10) Test10:Synonym

The tenth test suite is Synonym. The purpose is to test "Synonym". To complete this test, the corresponding selection in the query model should be chosen. And the parameters required by the Yahoo! Web Search APIs and keywords should be entered to generate the web search script. The "Topic Distillation" option is selected in the Query Model. The keyword is "SOA". The three parameters of the Yahoo! Search APIs are set to "query=SOA", "similar_ok=1" and "Results=100". Table 5.10 shows the specifications of the test suite..

Query Model: Synonym	Synonym
Query Model: Query Type	Topic distillation
Keyword	SOA
Yahoo! Search API parameter	query = SOA
Yahoo! Search API parameter	similar_ok = 1

Table 5.10: The Test Specifications of Synonym

Yahoo! Search API parameter	Results = 100
-----------------------------	---------------

5.1.2 Experimental Metrics

There are seven experimental metrics generated for each test as shown in table 5.11. "Response Time" is the time interval between entry of a request and receipt of a response. "Throughput" is the number of operations completed by the system per time unit. "Precision" is the percentage of retrieved data which is relevant. "Recall" is the percentage of relevant data which have been retrieved but here we use "Relative Recall" as explained in the chapter three. "Recall" and "Precision" are two important measures for evaluation of information retrieval. "TotalResultsAvailable" and "TotalResultsReturned" are two numbers that most web searches return. "Search Result" lists results returned by a web search service, including "Title", Time", "Summary" and "URL". "Response "Throughput", "TotalResultsAvailable", "TotalResultsReturned" and "Search Result" are automatically generated after each test is executed. We then click on every list from "Search Result" to check relevance and calculate "Precision" and "Relative Recall".

Experimental Metric	Description
Response Time	Interval between entry of a
	request and receipt of a response.
Throughput	Number of operations completed
	by the system per time unit.
Precision	The percentage of retrieved data
	those are relevant.
Relative Recall	The percentage of relevant data
	that have been retrieved.
TotalResultsAvailable	The number of query matches in
	the database.
TotalResultsReturned	The number of query matches
	returned.
Search Result	The list of query matches. Each
	includes Title, Summary and
	URL.
	• Title: the title of the web
	page.
	• Summary: summary text

Table 5.11: Experimental Metric

		associated	with	the	web
	•	page. URL: the U	JRL fo	or the	web
		page.			

5.2 Experimental Execution and Results

The ten baseline test cases have been executed with the research model and the prototype system. We have adopted Yahoo! Web Search APIs. In this section, we detail the results of each experiment with the selection of input parameters and the output test results.

5.2.1 Single Page- URL

In the first test, Single Page-Anchor Text, we click on "URL", "Topic Distillation" and input keyword "SOA". We click "Generate Script" as shown in Figure 5.1. There are two main result screens. Figure 5.2 shows the query script that is generated by the prototype sytem. The script illustrates the input requirements have been specified for the web search request. Figure 5.3 shows the search result of Test1. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.12 shows the metrics of the Test1 experiment. In addition, we perform a same query on Google and obtain the search result as shown in Figure 5.3 and metrics as shown in Table 5.13.

Generic Construct based Workload Model for Web Search		
d: SOA Generate Script		
Page Model		
 Anchor Text Font color Font size Frame Meta Table Title ✓ URL 		
For Example: Company Education Government Organization		
Query Model		
 Homepage finding Named page finding ✓ Topic Distillation 		
□ Authority-Hub □ PageRank		
 USM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency 		

Figure 5.1: Selected Input forTest1

Page Model:	Single Page-URL
Query Model:	Query Type-Topic Distillation
Keyword:	SOA
	hooapis.com/WebSearchService/V1/webSearch?

Figure 5.2: Script Output of Test1

Response time: 0.484 second Throughput: 2.06611570247934
totalResultsAvailable: 739000 totalResultsReturned: 100 Search Result :
1. <u>SOA - Society Of Actuaries</u> Education, research, and professional membership organization promoting its use in life insurance, health and retirement systems, and investments. http://www.soa.org/
2. <u>SOA - Society Of Actuaries</u> more about the results of an SOA - sponsored company survey on post - level premium 1 to participate in this unique SOA educational program, June 13-15 in http://www.soa.org/ccm/content/
3. <u>Service-Oriented Architecture</u> Oracle's Service Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology. http://www.oracle.com/technologies/soa/index.html
4. <u>Arch2Arch Online: Service-oriented Architecture</u> While these concepts have existed for decades, the adoption of SOA is BEA AquaLogic Service Bus 2.5: Integrating and Managing Services for SOA http://dev2dev.bea.com/soa/
5. <u>InformationWeek Software Soa Tech Center</u> http://www.soa-pipeline.com/
6. <u>SOA Resource Center, Service Oriented Architecture, SOA Appliance, SOA Architecture, System Integration</u> how BEA can help your organization successfully implement a SOA solution Understand the Fundamentals of SOA infrastructure and took SOA from pilot to http://www.bea.com/framework.jsp?CNT=index.htm&FP=/content/solutions/soa/
7. <u>All SOA Community Content on InfoQ</u> InfoQ.com (Information Queue) is an independent online community focused on change and innovation in enterprise software development, targeted primarily at the http://www.infoq.com/soa/
 創業 創業 調除網路

Figure 5.3: The Search Results of Test1

Web Images Video News Maps Gmail more ▼	Sign in 📤			
Google allinurt: soa Search Preferences				
Web Results 1 - 10 of about 915	5,000 for allinurl: soa. (0.11 seconds)			
SOA - Society Of Actuaries An educational, research and professional membership organization for actuaries in the US and Canada who primarily practice in pension, employee benefits, www.soa.org/ - 32k - Cached - Similar pages webservices.xml.com: What Is Service-Oriented Architecture Service-Oriented Architecture underpins most modern web services. It aims to achieve loose coupling between interacting software agents in order to preserve webservices.xml.com/pub/a/ws/2003/09/30/soa.html - 80k - Cached - Similar pages SOA Software - SOA, XML, and Web services management, SOA SOA Management, Security, and Governance Company. Provides software and solutions	Sponsored Links Soa 提供Sun® 培訓及專業認證 現在就加入 Sun.com! tw.sun.com Calculate SOA benefits Analyze the value of SOA for your organization! www.soavalueassessment.com Open Source SOA Platform			
for Web Services Management, Security, and Governance, (a) Map of 1011 Dry Creek Dr. Longmont, CO 80503 www.soa.com/- 51k - 9 Jul 2007 - <u>Cached</u> - <u>Similar pages</u>	LogicBlaze FUSE - run SOA now Free white paper, product download open.iona.com/			
Service-Oriented Architecture Oracle's Service Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology. www.oracle.com/technologies/soa/index.html - 52k - <u>Cached</u> - <u>Similar pages</u>	SOA Portal Architecture, Services, Solutions, Web Services Directory, News, Jobs www.soahub.com			
What is SOA, really? What is SOA, really? 13. Posted by Uncle Bob Wed, 11 Apr 2007 12:17:44 GMT. The good news is, you probably already know. The bad news is, you probably know typo.objectmentor.com/articles/2007/04/11/what.is-soa-really - 26k - <u>Cached</u> - <u>Similar pages</u> SOA WORLD MAGAZINE				
Jump to a SYS-CON Magazine .NET Developer's JournalNETDJ, AjaxWorld Magazine - AWM. IBM Product Showcase. ColdFusion Developer's Journal - CFDJ				
 (introduct of lowed as cooling and Developer a bounding of Data (introduct of lowed as cooling and lowed as cooling as cooling and lowed as cooling and lowed as cooling and lowed as cooling as cooling				

Figure 5.4: The Search Results of Test1 by Google

1.	Response Time	0.484 second
2.	Throughput	2.06611570247934
3.	Precision	0.5
4.	Relative Recall	0.48
5.	TotalResultsAvailable	712000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.3

Table 5.12: Experimental Metrics of Test1

Table 5.13: Experimental Metrics of Test1 by Google

1.	Response Time	0.11 second
2.	Throughput	9.09090909090909
3.	Precision	0.6
4.	Relative Recall	0.42
5.	TotalResultsAvailable	915000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.4

5.2.2 Single Page- Font Size, Font Color, Frame, Meta and Table

In the second test, the Single Page- Font Size, Font Color, Frame, Meta and Table, we click on "Font Size", "Font Color", "Frame", "Meta", "Table" and "Topic Distillation" and input keyword "SOA". We click on "Generate Script" as shown in Figure 5.5. There are two main result screens. Figure 5.6 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified for the web search request. Figure 5.7 shows the search result of Test2. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.14 shows the metrics of the test2. In addition, we perform a same query on Google and obtain the search result as shown in Figure 5.8 and metrics as shown in Table 5.15.

\mathbf{A}	Generic Construct based Workload Model for Web Search	-
Keyword: SOA	Generate Script	
-	Page Model	
Single-Page	 Anchor Text ✓ Font color ✓ Font size ✓ Frame ✓ Meta ✓ Table Title 	
Multi-Page	For Example: Company Education Government Organization	
	Query Model	
QueryType	 ☐ Homepage finding ☐ Named page finding ☑ Topic Distillation 	
Link Structure	□ Authority-Hub □ PageRank	
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency 	
Synonym	Synonym	

Figure 5.5: Selected Input for Test2

Page Model:	Single Page-Font color, Font color, Font color, Font color, Font color
Query Model:	Query Type-Topic Distillation
Keyword:	SOA
	hooapis.com/WebSearchService/V1/webSearch? IWXpK9430f27iIrYy8.wB429NQLAgoz8MqZDrYjFSsk6Y6icZCZ7I- ry=SOA

Figure 5.6: Script Output of Test2

Response time: 0.765 second Throughput: 1.30718954248366 totalResultsAvailable: 32800000
totalResultsReturned: 100
Search Result :
1.SOA - Society Of Actuaries
Education, research, and professional membership organization promoting its use in life insurance, health and retirement systems,
and investments.
http://www.soa.org/
mps, an about of
2. Service-oriented architecture - Wikipedia, the free encyclopedia
SOA can also be regarded as a style of information systems architecture that SOA-based systems can therefore be independent of
development technologies and
http://en.wikipedia.org/wiki/Service-oriented_architecture
indp://en.wikipedia.org/wiki/service-onented_arcintecture
3.SOA Watch
Seeks to close the U.S. Army School of the Americas through vigils and fasts, demonstrations and nonviolent protest, as well as
media and legislative work.
http://www.soaw.org/
mp.//www.soaw.org/
4. Service-Oriented Architecture
4. Service-Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology.
http://www.oracle.com/technologies/soa/index.html
5.SOA - Society Of Actuaries
Careers at SOA. Course Catalog. Exams. Resources. Continuing Education 1 to participate in this unique SOA educational
program, June 13-15 in Seattle
http://www.soa.org/ccm/content
ings/www.soa.org/com/com/cm
6.Arch2Arch Online: Service-oriented Architecture
While these concepts have existed for decades, the adoption of SOA is BEA AquaLogic Service Bus 2.5: Integrating and
Manaing Services for SOA
 ◎ 完成

Figure 5.7: The Search Results of Test2

Web Images Video News Maps Gmail more ▼	Sign in			
Google allintext: soa Search Advanced Ser	arch			
Web Results 1 -	10 of about 77,400,000 for allintext: soa. (0.17 seconds)			
Service-oriented architecture - Wikipedia, the free encyclop There is no widely-agreed upon definition of service-oriented architecture other tha literal translation that it is an architecture that relies on en.wikipedia.org/wiki/Service-oriented_architecture - 93k - <u>Cached</u> - <u>Similar pages</u> <u>SOA - Wikipedia, the free encyclopedia</u> SOA stands for Service Oriented Architecture - exposing critical Business				
Services/applications to the world through Selfdescribing WebServices en wikipedia.org/wiki/SOA - 18k - <u>Cached</u> - <u>Similar pages</u> SOA - Society Of Actuaries	Calculate the benefits of SOA for your organization! www. soa valueassessment.com			
An educational, research and professional membership organization for actuaries in t and Canada who primarily practice in pension, employee benefits, www.soa.org/ - 32k - <u>Cached</u> - <u>Similar pages</u>	LogicBlaze FUSE - run SOA now Free white paper, product download open.iona.com/			
webservices.xml.com: What Is Service-Oriented Architectur Service-Oriented Architecture underpins most modern web services. It aims to ach loose coupling between interacting software agents in order to preserve webservices.xml.com/pub/a/ws/2003/09/30/soa.html - 80k - Cached - Similar pages				
SOA Software - SOA, XML, and Web services management, SOA SOA Management, Security, and Governance Company. Provides software and solutions for Web Services Management, Security, and Governance, @Map of 1011 Dry Creek Dr. Longmont, CO 80503 www.soa.com/- 51k - 9 Jul 2007 - <u>Cached</u> - <u>Similar pages</u>				
IBM Service Oriented Architecture (SOA) A service-oriented architecture (SOA) provides an application framework that turns business applications into individual business functions and processes,	5 🖉 網際網路			
1				

Figure 5.8: The Search Results of Test2 by Google

1.	Response Time	0.765
2.	Throughput	1.30718954248366
3.	Precision	0.52
4.	Recall	0.605
5.	TotalResultsAvailable	32800000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.6

Table 5.14: Experimental Metrics of test2

Table 5.15: Experimental Metrics of test2 by Google

8.	Response Time	0.17 second
9.	Throughput	5. 882352941176
10.	Precision	0.53
11.	Recall	0.61
12.	TotalResultsAvailable	77400000
13.	TotalResultsReturned	100
14.	The Search Result	As show in Fig 5.8

5.2.3 Single Page- Title

In the third test, the Single Page- Title, we click on "Title" and" Topic Distillation" and input keyword "SOA". We click on "Generate Script" as shown in Figure 5.9. There are two main result screens. Figure 5.10 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified for the web search request. Figure 5.11 shows the search result of Test1. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.16 shows the metrics of the test3. In addition, we perform a same query on Google and obtain the search result as shown in Figure 5.12 and metrics as shown in Table 5.17.

Kevword: SOA Generate Script		
Page Model		
Single-Page	Anchor Text Font color Font size Trame Meta Table V Title	
Multi-Page	For Example: Company Education Government Organization	
	Query Model	
QueryType	 ☐ Homepage finding ☐ Named page finding ☑ Topic Distillation 	
Link Structure	□ Authority-Hub □ PageRank	
Similarity	 USM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency 	
Synonym	Synonym	

Figure 5.9: Selected Input for Test3

Page Model:	Single Page-Title
Query Model:	Query Type-Topic Distillation
Keyword:	SOA
	HwXpK9430fZ7iIrYy8.wB429NQLAgoz8MqZDrYjFSsk6Y6icZCZ7I- ry=intitle:SOA
results=100&que	

Figure 5.10: Script Output of Test3

Response time: 0.359 second Throughput: 2.78551532033426 totalResultsAvailable: 713000 totalResultsReturned: 100 Search Result :
1. <u>SOA - Society Of Actuaries</u> Education, research, and professional membership organization promoting its use in life insurance, health and retirement systems, and investments. http://www.soa.org/
2. <u>SOA Watch</u> Seeks to close the U.S. Army School of the Americas through vigils and fasts, demonstrations and nonviolent protest, as well as media and legislative work. http://www.soaw.org/
3. <u>SOA - Society Of Actuaries</u> Careers at SOA. Course Catalog. Exams. Resources. Continuing Education 1 to participate in this unique SOA educational program, June 13-15 in Seattle http://www.soa.org/ccm/content
4. <u>SOA Resource Center, Service Oriented Architecture, SOA Appliance, SOA Architecture, System Integration</u> how BEA can help your organization successfully implement a SOA solution Understand the Fundamentals of SOA infrastructure and took SOA from pilot to http://www.bea.com/framework.jsp?CNT=index.htm&FP=/content/solutions/soa/
5. <u>SOA Community on InfoQ</u> InfoQ.com (Information Queue) is an independent online community focused on change and innovation in enterprise software development, targeted primarily at the http://www.infoq.com/soa/
6. <u>SOA - CIO.com - Business Technology Leadership</u> Everything you need to know about the advantages, disadvantages and challenges of service oriented architecture The Past,

Figure 5.11: The Search Results of Test3

Web Images <u>Video</u> <u>News</u> <u>Maps</u> <u>Gmail</u> more ▼	Sign in
Google allintitle: soa Search Preferences	
Web Results 1 - 10 of about 847,0	000 for allintitle: soa. (0.08 seconds)
SOA - Society Of Actuaries An educational, research and professional membership organization for actuaries in the US and Canada who primarily practice in pension, employee benefits, www.soa.org/ - 32k - Cached - Similar pages SOA Software - SOA, XML, and Web services management, SOA SOA Management, Security, and Governance Company. Provides software and solutions for Web Services Management, Security, and Governance, ®Map of 1011 Dry Creek Dr. Longmont, CO 80503 www.soa.com/ - 51k - 9 Jul 2007 - Cached - Similar pages	Sponsored Links Soa Professional IT Courses Find out more! tw.sun.com SOA Value Assessment Calculate the benefits of SOA for your organization! www.soavalueassessment.com
SOA - Wikipedia, the free encyclopedia From Wikipedia, the free encyclopedia. Jump to: navigation, search. SOA stands for Service Oriented Architecture - exposing critical Business en.wikipedia.org/wiki/ SOA - 18k - <u>Cached</u> - <u>Similar pages</u>	Create AJAX and SOA Apps Use NXJ Developer to quickly create Rich Internet and SOA applications www.unify.com
IBM Service Oriented Architecture (SOA) A service-oriented architecture (SOA) provides an application framework that turns business applications into individual business functions and processes, www.ibm.com/soa - 38k - <u>Cached</u> - <u>Similar pages</u>	Open Source SOA Platform LogicBlaze FUSE - run SOA now Free white paper, product download open.iona.com/
Service-oriented architecture (SOA) definition The definition of a service-oriented architecture (soa) involving services and connections (includes graphic). www.service-architecture.com/web-services/articles/service- oriented_architecture_soa_definition.html - 31k - 8 Jul 2007 - <u>Cached</u> - <u>Similar pages</u>	SOA Portal Architecture, Services, Solutions, Web Services Directory, News, Jobs www.soahub.com
SOA Watch The US Army School of Americas (SOA), based in Fort Benning, Georgia, trains Latin American security personnel in combat, counter insurgency 劉完成	▼ ● 新院網路

Figure 5.12: The Search Results of Test3 by Google

1.	Response Time	0.359 second
2.	Throughput	2.78551532033426
3.	Precision	0.6
4.	Recall	0.6
5.	TotalResultsAvailable	713000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.11

Table 5.16: Experimental Metrics of test3

Table 5.17: Experimental Metrics of test3 by Google

8.	Response Time	0.08 second
9.	Throughput	12.5
10.	Precision	0.61
11.	Recall	0.59
12.	TotalResultsAvailable	847000
13.	TotalResultsReturned	100
14.	The Search Result	As show in Fig 5.12

5.2.4 Multi-Page-Company

In the fourth test, the Multi-Page-Company, we click on "Company" and "Topic Distillation" and input keyword "SOA". We click on "Generate Script" as shown in Figure 5.13. There are two main result screens. Figure 5.14 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified for the web search request. Figure 5.15 shows the search result of Test4. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.18 shows the metrics of the test4.

A Generic Construct based Workload Model for Web Search		
Keyword: SOA Generate Script		
Page Model		
Single-Page	 Anchor Text Font color Font size Frame Meta Table Title 	
Multi-Page	For Example: Company Education Government Organization ibm	
	Query Model	
QueryType	 □ Homepage finding □ Named page finding ☑ Topic Distillation 	
Link Structure	□ Authority-Hub □ PageRank	
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency 	
Synonym	Synonym	
	S) 近端内部網路	

Figure 5.13: Selected Input for Test4

Page Model:	Multi-Page-Company,www.ibm.com
Query Model:	Query Type-Topic Distillation
Keyword:	SOA
appid=YaWphs_V34	hooapis.com/WebSearchService/V1/webSearch? A HwXpK9430f27iIrYy8.wB429NQLAgoz8HqZDrYjFSsk6Y6icZCZ7I- e=www.ibm.com&query=SOA
Submit Reset	

Figure 5.14 Script Output of Test4

完成 🧼 網際網路
ompany to more closely align and
IBM SOA Governance, service oriented architecture, SOA OA governance enables organizations to realize the full potential of Service An SOA is a business strategy that enables a
tp://www.ibm.com/developerworks/webservices/library/ws-antipatterns/
. Service-Oriented Architecture (SOA) antipatterns, which are descriptions successful implementations of SOA are becoming ore SOA: brief introduction
<u>SOA antipatterns</u>
tp://www.ibm.com/developerworks/webservices/newto/
want to find out more about what SOA is and how it can help your organization, SOA entry points Using SOA you can, for sample, create service-based
New to SOA and Web services
ttp://www.ibm.com/developerworks/webservices/library/ws-soaintro.html
leveloperWorks, April 2004)
oday's Web services implementations are typically simple and often similar to a for more information and resources on SOA
Service-Oriented Architecture expands the vision of Web services, Part 1
ttp://www.ibm.com/developerworks/webservices
or creating Web services and
he SOA and Web services section on the developerWorks Web site is your resource for tools, demos, articles, tips, and standards
developerWorks : SOA and Web services
ttp://www.ibm.com/soa
BM has proven experience, software, and services to help implement an SOA
IBM Service Oriented Architecture (SOA)
earch Result :
talResultsReturned: 100
talResultsAvailable: 217000
hroughput: 8
esponse time: 0.125 second hroughput: 8

Figure 5.15: The Search Results of Test4

	Table 5.16. Experimental Metries of test-		
1.	Response Time	0.125 second	
2.	Throughput	8	
3.	Precision	0.9	
4.	Recall	0.8	
5.	TotalResultsAvailable	217000	
6.	TotalResultsReturned	100	
7.	The Search Result	As show in Fig 5.12	

Table 5.18: Experimental Metrics of test4

5.2.5 Query Type-Homepage Finding

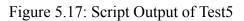
In the fifth test, the Query Type-Homepage finding, we click on "Homepage Finding" and input keyword "ibm". We click on "Generate Script" as shown in Figure 5.16. There are two main result screens. Figure 5.17 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified for the web search request Figure 5.18 shows the search result of Test5. "TotalResultsAvailable", "TotalResultsReturned" and the lists from

"Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.19 shows the metrics of the test5.

Keyword: ibm	Generate Script
	Page Model
Single-Page	 Anchor Text Font color Font size Frame Meta Table Title
Multi-Page	For Example: © Company □ Education □ Government □ Organization
	Query Model
QueryType	 ☑ Homepage finding □ Named page finding □ Topic Distillation
Link Structure	□ Authority-Hub □ PageRank
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency
Synonym	Synonym

Figure 5.16: Selected Input for Test5

Page Model:	Multi-Page-Company
Query Model:	Query Type-Homepage finding
Keyword:	ibm
	hooapis.com/WebSearchService/V1/webSearch?



Response time: 0.578 second Throughput: 1.73010380622837 totalResultsAvailable: 472000 totalResultsReturned: 100 Search Result :
1.IBM 中国 The IBM corporate home page, entry point to information about IBM products and services http://www.ibm.com/cn/zh
2. <u>IBM Intellectual Property And Licensing</u> Overview Intellectual Property & Licensing, patent portfolio, patent licensing practices, technology licensing practices, IBM Technical Disclosure Bulletin http://www.ibm.com/patents
3. <u>IBM Slovensko</u> IBM Slovensko a jej domovská stránka http://www.ibm.com/sk
4. <u>IBM Österreich</u> IBM in Österreich bietet Ihnen ein umfangreiches Angebot von IT-Produkten, -Dienstleistungen und -Lösungen für Ihr Unternehmen. http://www.ibm.com/at
5. <u>IBM Support & download - Danmark</u> Velkommen til ibm.com's support portal. Fra denne side kan du søge i vores videndatabase for teknisk support, hente drivere og downloade software, få adgang til http://www.ibm.com/support/dk/
6. <u>developerWorks : IBM's resource for developers</u> Access tools, code, training, forums, standards, and how-to documentation for Rational, WebSphere, DB2, Lotus, Tivoli, Power Architecture, and e-Server, plus open http://www.ibm.com/developerworks/
高)完成

Figure 5.18: The Search Results of Test5

1.	Response Time	0.578 second
2.	Throughput	1.73010380622837
3.	Precision	1
4.	Recall	1
5.	TotalResultsAvailable	447200
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.15

Table 5.19: Experimental Metrics of test5

5.2.6 Query Type- Named page Finding

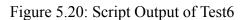
In the sixth test, the Query Type- Named page Finding, we click on "Named page Finding" and input keyword "help.yahoo.com/help/us/ysearch/tips/tips-04.html". We click on "Generate Script" as shown in Figure 5.19. There are two main result screens. Figure 5.20 shows the query script that is generated by the web search service. The script illustrates the input requirements have been specified for the web search request. Figure 5.21 shows the search result of Test6.

"TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.20 shows the metrics of the test6.

Kevword: help.yahoo.c	com/help/us/ysearch/tips/tips-04.html	
	Page Model	
Single-Page	Anchor Text Font color Font size Meta Table Title	
Multi-Page	For Example: Company Education Government Organization	
	Query Model	
QueryType	 ☐ Homepage finding ☑ Named page finding □ Topic Distillation 	
Link Structure	□ Authority-Hub □ PageRank	
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency 	
Synonym	Synonym	

Figure 5.19: Selected Input for Test6

Page Model:	
Query Model:	Query Type-Named page finding
Keyword:	help.yahoo.com/help/us/ysearch/tips/tips-04.html
ttp://search.ya	



Response time: 0.718 second Throughput: 1.39275766016713 totalResultsAvailable: 1 totalResultsReturned: 1 Search Result :	
1. <u>Yahoo! Help - Search</u> Search finds what you tell it to, so here are some tips for better results Use specific words to describe exactly what you're looking for http://help.yahoo.com/help/us/ysearch/tips/tips-04.html	
 完成 	

Figure 5.21: The Search Results of Test6

	1	
1.	Response Time	0.718 second
2.	Throughput	1.39275766016713
3.	Precision	1
4.	Recall	1
5.	TotalResultsAvailable	1
6.	TotalResultsReturned	1
7.	The Search Result	As show in Fig 5.18

5.2.7 Query Type-Topic distillation

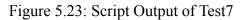
In the seventh test, Query Type-Topic distillation, we click on "Topic Distillation" and input keyword "ibm". We click on "Generate Script" as shown in Figure 5.22. There are two main result screens. Figure 5.23 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified for the web search request. Figure 5.24

shows the search result of Test7. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.21 shows the metrics of the test7.

Keyword: SOA	Generate Script
	Page Model
Single-Page	 Anchor Text Font color Font size Frame Meta Table Title
Multi-Page	For Example: Company Education Government Organization
	Query Model
QueryType	 □ Homepage finding □ Named page finding ☑ Topic Distillation
Link Structure	□ Authority-Hub □ PageRank
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency
Synonym	Synonym

Figure 5.22: Selected Input for Test7

Page Model:	
Query Model:	Query Type-Topic Distillation
Keyword:	SOA
ppid=YaWphs_V34	hooapis.com/WebSearchService/V1/webSearch? MwXpK9430f277IrVy8.wB429NQLAgoz8MqZDrVjFSsk6Y6icZCZ7I- ry=SOA



Response time: 0.359 second Throughput: 2.78551532033426 totaIResultsAvailable: 33000000 totaIResultsReturned: 100 Search Result :
1.SOA - Society Of Actuaries Education, research, and professional membership organization promoting its use in life insurance, health and retirement systems, and investments. http://www.soa.org/
2. <u>Service-oriented architecture - Wikipedia, the free encyclopedia</u> Service-oriented architecture (SOA [pronounced "sō-uh" or "es-ō-ā"]) describes a software architecture that defines the use of loosely coupled software services http://en.wikipedia.org/wiki/Service-oriented_architecture
3. <u>SOA Watch</u> Seeks to close the U.S. Army School of the Americas through vigils and fasts, demonstrations and nonviolent protest, as well as media and legislative work. http://www.soaw.org/
4. <u>Service-Oriented Architecture</u> Oracle's Service Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology. http://www.oracle.com/technologies/soa/index.html
5. <u>SOA - Society Of Actuaries</u> Careers at SOA. Course Catalog. Exams. Resources. Continuing Education 1 to participate in this unique SOA educational program, June 13-15 in Seattle http://www.soa.org/ccm/content
6. <u>Arch2Arch Online: Service-oriented Architecture</u> While these concepts have existed for decades, the adoption of SOA is BEA AquaLogic Service Bus 2.5: Integrating and Managing Services for SOA

Figure 5.24: The Search Results of Test7

	ruble 5.21. Experimental Weth	
1.	Response Time	0.359 second
2.	Throughput	2.78551532033426
3.	Precision	0.5
4.	Recall	0.45
5.	TotalResultsAvailable	33000000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.21

Table 5.21: Experimental Metrics of test7

5.2.8 Link Structure-Authority-Hub

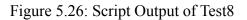
In the eighth test, Link Structure-Authority-Hub, we click on "Authority-Hub" and" Multi-Page-Company" and input keyword "soa" and "ibm". We click on "Generate Script" as shown in Figure 5.25. There are two main result screens. Figure 5.26 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified

for the web search request. Figure 5.27 shows the search result of Test8. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.22 shows the metrics of the test8.

Keyword: SOA	Generate Script
	Page Model
Single-Page	 Anchor Text Font color Font size Frame Meta Table Title
Multi-Page	For Example: Company Education Government Organization ibm
	Query Model
QueryType	 Homepage finding Named page finding Topic Distillation
Link Structure	✓Authority-Hub□ PageRank
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency
Synonym	Synonym

Figure 5.25: Selected Input for Test8

Page Model:	Multi-Page-Company,www.ibm.com
Query Model:	Authority-Hub
Keyword:	SOA
appid=YaWphs_V34	hooapis.com/WebSearchService/V1/webSearch? MwXpK9430f27iIrYy8.wB429NQLAgoz8HqZDrYjFSsk6Y6icZCZ7I- ry=link:http://www.ibm.com+SOA



Response time: 0.406 second Throughput: 2.46305418719212 totalResultsAvailable: 5080000 totalResultsReturned: 100 Search Result :	
1. <u>Service-oriented architecture - Wikipedia, the free encyclopedia</u> Service-oriented architecture (SOA [pronounced "sō-uh" or "es-ō-ā"]) describes a Service-Oriented Architecture Implementation Framework. Service Oriented http://en.wikipedia.org/wiki/Service-oriented_architecture	
2. Web Services and Service-Oriented Architectures and services that can be used to develop a service-oriented architecture covers the type of service-oriented architecture and related technologies shown http://www.service-architecture.com/index.html	
3. <u>Service Oriented Architecture</u> Learn About Service Oriented Architecture intelligence (BI) and service-oriented architecture (SOA) have conflicting principles and needs http://msdn2.microsoft.com/en-us/architecture/aa948857.aspx	
4. <u>Service-oriented architecture (SOA) definition</u> The definition of a service-oriented architecture (soa) involving services and Architecture and Service-Oriented Architecture, and helps organizations http://www.service-architecture.com/web-services/articles/service-oriented_architecture_soa_definition.html	
5. <u>Service Oriented Architecture</u> The Service Oriented Architecture (SOA) is a distributed software model Web Services are considered an example of a Service Oriented Architecture http://www.serviceoriented.org/service_oriented_architecture.html	
 6. <u>Service-Oriented Architecture</u> Oracle's Service Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology. ● 網際網路 	•

Figure 5.27: The Search Results of Test8

	Tuble 5.22. Experimental Met	
1.	Response Time	0.406 second
2.	Throughput	2.46305418719212
3.	Precision	0.72
4.	Recall	0.7
5.	TotalResultsAvailable	5080000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.24

Table 5.22: Experimental Metrics of Test8

5.2.9 Similarity-TLS

In the ninth test, Similarity-TLS, we click on "TLS" and" Topic Distillation" and input keyword "Service Oriented Architecture". We click on "Generate Script" as shown in Figure 5.28. There are two main result screens. Figure 5.29 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified for the web search request. Figure 5.30 shows the search result of Test9. "TotalResultsAvailable", "TotalResultsReturned" and

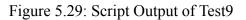
the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.23 shows the metrics of the test9.

.

1	Constants Societ
Keyword: Service Orier	
	Page Model
Single-Page	 Anchor Text Font color Font size Frame Meta Table Title
Multi-Page	For Example: Company Education Government Organization
	Query Model
QueryType	Homepage finding Named page finding Topic Distillation
Link Structure	□ Authority-Hub □ PageRank
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency
Synonym	

Figure 5.28: Selected Input for Test9

Page Model:	
Query Model:	Query Type-Topic Distillation, Similarity-TLS (Three-Level Scoring Method): sub-phrase frequency
Keyword:	Service Oriented Architecture
	hooapis.com/WebSearchService/V1/webSearch?



Response time: 0.343 second Throughput: 2.91545189504373 totalResultsAvailable: 33500000 totalResultsReturned: 100 Search Result :
1. <u>SOA - Society Of Actuaries</u> Education, research, and professional membership organization promoting its use in life insurance, health and retirement systems, and investments. http://www.soa.org/
2. <u>Service-oriented architecture - Wikipedia</u> , the free encyclopedia Service-oriented architecture (SOA [pronounced "sō-uh" or "es-ō-ā"]) describes a software architecture that defines the use of loosely coupled software services http://en.wikipedia.org/wiki/Service-oriented_architecture
3. <u>SOA Watch</u> Seeks to close the U.S. Army School of the Americas through vigils and fasts, demonstrations and nonviolent protest, as well as media and legislative work. http://www.soaw.org/
4. <u>Service-Oriented Architecture</u> Oracle's Service Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology. http://www.oracle.com/technologies/soa/index.html
5. <u>SOA - Society Of Actuaries</u> Careers at SOA. Course Catalog. Exams. Resources. Continuing Education 1 to participate in this unique SOA educational program, June 13-15 in Seattle http://www.soa.org/ccm/content
6. <u>Arch2Arch Online: Service-oriented Architecture</u> While these concepts have existed for decades, the adoption of SOA is BEA AquaLogic Service Bus 2.5: Integrating and Managing Services for SOA

Figure 5.30: The Search Results of Test9

	Tuble 5.25. Experimental Wet	
1.	Response Time	0.343 second
2.	Throughput	2.91545189504373
3.	Precision	0.735
4.	Recall	0.825
5.	TotalResultsAvailable	33500000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.27

Table 5.23: Experimental Metrics of Test9

5.2.10 Synonym

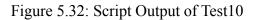
In the tenth test, Similarity-TLS we click on "Synonym" and" Topic Distillation" and input keyword "SOA". We click on "Generate Script" as shown in Figure 5.31. There are two main result screens. Figure 5.32 shows the query script that is generated by the prototype system. The script illustrates the input requirements have been specified foro the web search request. Figure

5.33 shows the search result of Test10. "TotalResultsAvailable", "TotalResultsReturned" and the lists from "Search Result" are returned by the Yahoo! web search services. "TotalResultsAvailable" is set with "Response Time" and "Throughput" as the main metrics to be automatically generated. Based on the search results, we will check the relevance of each URL and calculate precision and recall. Table 5.24 shows the metrics of the test10.

Keyword: SOA	Generate Script
	Page Model
Single-Page	Anchor Text Font color Font size Frame Meta Table Title
Multi-Page	For Example: Company Education Government Organization
	Query Model
QueryType	 ☐ Homepage finding ☐ Named page finding ☑ Topic Distillation
Link Structure	□ Authority-Hub □ PageRank
Similarity	 VSM(Vector Space Model): term frequency Okapi(Okapi Measurement Method): term frequency and document length CDR(Cover Density Ranking): the position of term TLS(Three-Level Scoring Method): sub-phrase frequency
Synonym	✓ Synonym

Figure 5.31: Selected Input for Test10

Page Model:	
Query Model:	Query Type-Topic Distillation,Synonym-Synonym
Keyword:	SOA
appid=YaWphs_V34	hooapis.com/WebSearchService/V1/webSearch?



Response time: 0.531 second Throughput: 1.88323917137476 totalResultsAvailable: 33500000 totalResultsReturned: 100 Search Result :
1. <u>SOA - Society Of Actuaries</u> Education, research, and professional membership organization promoting its use in life insurance, health and retirement systems, and investments. http://www.soa.org/
2. <u>Service-oriented architecture - Wikipedia, the free encyclopedia</u> Service-oriented architecture (SOA [pronounced "sō-uh" or "es-ō-ā"]) describes a software architecture that defines the use of loosely coupled software services http://en.wikipedia.org/wiki/Service-oriented_architecture
3. <u>SOA - Society Of Actuaries</u> ASSA Honors SOA Past President. May 31, 2007 1 to participate in this unique SOA educational program, June 13-15 in Seattle. http://www.soa.org/ccm/content/
4. <u>SOA Watch</u> Seeks to close the U.S. Army School of the Americas through vigils and fasts, demonstrations and nonviolent protest, as well as media and legislative work. http://www.soaw.org/
5. <u>Service-Oriented Architecture</u> Oracle's Service Oriented Architecture (SOA) solutions. Create, deploy, and manage SOAs with best-of-breed Oracle technology. http://www.oracle.com/technologies/soa/index.html
6. <u>Arch2Arch Online: Service-oriented Architecture</u> While these concepts have existed for decades, the adoption of SOA is BEA AquaLogic Service Bus 2.5: Integrating and Managing Services for SOA

Figure 5.33: The Search Results of Test10

P		
1.	Response Time	0.531 second
2.	Throughput	1.88323917137476
3.	Precision	0.375
4.	Recall	0.4
5.	TotalResultsAvailable	33500000
6.	TotalResultsReturned	100
7.	The Search Result	As show in Fig 5.30

Table 5.24: Experimental Metrics of Test10

5.3 Summary

Some uncontrollable factors can affect the search results on the internet. It would be best to execute all tests simultaneously in sequence because the internet is always changing and growing. The ten test suites are finished by requesting Yahoo! Web Search Services so the tests are executed in a web environment. Web pages are always increasing and the status of network traffic can affect the efficiency of the web search so the relevance assessments must be carried out as soon as the hit lists have been retrieved. Otherwise, the differences in performance metrics between test cases will be meaningless. The performance metrics "Response Time" and "Throughput" are shorter during rush hours. The value of the "TotalResultsAvailable" metric is greater than the search results generated few days previously. In order to maintain consistency of the entire experiment, it is better to execute each test immediately in sequence.

Each test suite provides a different workload to benchmark the web search. And we input the same keyword "SOA" and "Service Oriented Architecture" in each test. Consequently, the ranking of each search result is somehow different. From this, we derive some conclusions on the different rankings. Because a web search is a kind of text search, a page that matches the keyword exactly might not match what users want to find. For example, in the third test" Single Page-Title", the top three results match "SOA" but not "Service Oriented Architecture". In addition, the eighth test "Link Structure-Authority-Hub" performs a text search and searches pages linked to the URL www.ibm.com. This can find pages related to "Service Oriented Architecture" as well as "SOA" and the accuracy of the top five pages from the search result is better than that for the other tests. Obviously, this is the reason why designers of web search engines are always trying to refine ranking algorithms.

An artificial calculation of precision and recall cannot be error free in every case. We need to click on every list of search results to verify whether a page is relevant or not. Although we follow the relevance criteria in the literature, in some cases correct items may be missed or misjudged as artificially working. There are two ways to limit errors. One way is to use the judgments of more than two people and get an average. The other way is to choose the more reliable keyword because it makes judgment of a page's relevance easier.