Location Oriented Integration of Internet Information - Mobile Info Search -

Katsumi Takahashi[1], Seiji Yokoji[1], and Nobuyuki Miura[2]

[1] NTT Information Sharing Platform Laboratories,
Nippon Telegraph and Telephone Corporation
3-9-11, Midori-cho, Musashino-Shi, Tokyo, 180-8585 Japan
takahashi.katsumi@lab.ntt.co.jp, yokoji.seiji@lab.ntt.co.jp

[2] NTT DoCoMo Multimedia Laboratories,
3-5, Hikarinooka, Yokosuka, Kanagawa, 239-8536, Japan
miura@mml.yrp.nttdocomo.co.jp

Abstract. Information on the Internet is becoming more attractive and useful for our daily life. It provides things on the town, happenings on the city, and learning of the real world. If we can utilize such information for the interaction between the human and the city, it can enhance the value and the function of the city. In this paper we introduce the research project "*Mobile Info Search*" in which we study the method of integrating heterogeneous information in a location-oriented way for providing it in a handy form with mobile computing. We have a prototype of *Mobile Info Search* at *http://www.kokono.net/*, a location-based "search engine". Local information such as yellow pages, maps, and relevant Web pages at any location of Japan are provided with a simple interface. From the analysis of test services, we will discuss the user issues and information source issues; What kind of local information is welcomed? What can we learn from collected documents? Through the experience of handling various contents related to the real-world, we describe the potential of the Internet information for the digital city efforts.

1 Introduction

Not only for researches and businesses, information on the Internet is becoming useful for our daily life today. There is much attractive information available. It includes restaurant guides, local maps, public transportation, and weather reports. Moreover, due to the progress of the mobile computing, we can access to the Internet even when we are not in offices or home. We can use Internet just at the city. We think the function of the city is to provide the information and the marketplace for any activities. Though we can get rich information directly from the city, the Internet information from the each constituents of the city has a potential to enhance the value and the function of the city. That what we think about the digital city. A methodology and a service for utilizing such information from the Internet is required for our coming digital city.

This research is to study the computing methodology of overlapping the real world and the information from the network. If we can apply these ideas to the interaction between the real world and us, we can know more about the world and the city. If we use information on the Internet before we go to the city, we will be able to have more chance to find the city deeply. If we use it just on the city using mobile computing, we can know more about the things on the city. For these targets, one goal is to compile the "personalized digital guidebook" automatically from the Internet resources. Commercially there can be a portal site (the server which provides the index to enter the Internet) services for mobile users. And furthermore, not only using but also supplying and sending local information from each user to others relevant to the location, the open space for sharing the local information can be created.

Mobile Info Search is the name of our research project and an Internet based application[1]. The goal of *Mobile Info Search* is to study the location-oriented computing. We now concentrate to integrate local information existing on the Internet into a handy form especially for mobile-computing users. *Mobile Info Search* is characterized by the usage of "location information". Location information represents the geographical position or the area of the information in the form of address strings, longitude-latitude, landmarks,... etc. *Mobile Info Search* uses the location information of Internet resources for information integration. So the location of the mobile user and the location handling such as extraction and transformation of the information source play important role. We call this integration "location-oriented information integration".

Mobile Info Search was implemented experimentally on the Internet in 1997 and has been open to the public since then. Our prototype *Mobile Info Search* service is at *http://www.kokono.net/*[1]. It provides local information ("kokono" information in Japanese) about shops, maps, the weather, transportation, etc.

In this paper we first describe an outline of *Mobile Info Search* in section 2, and two features; *Location-Oriented Meta Search* and *Location-Oriented Robot-Based Search* in section 3 and 4. In section 5 and 6 we discuss the results from the experiment. After a look at related work, we conclude with a summary.

2 Mobile Info Search

2.1 Local information on the Internet

On the Internet, we can find information related to a certain location. It describes about shops, towns, and sights on the location or about the region itself. We call it "local information" in this article. *Mobile Info Search* uses such resources existing on the Internet.

Most local information today is divided into two types by their stored forms. Database type and static file type. Examples of database-type resources provided on the Internet are shown in Table. 1. They are provided through the CGI program of the WWW server. The access method and the use of location

[1] see *http://www.kokono.net/english/* for English

information is vary for each services. A mediation service for these resources is described in the section 3 of *Location-Oriented Meta Search*.

Table 1. Database-type resources on the Internet

Services	Location information used for the search
Maps	longitude-latitude
Yellow Pages	address (and categories,... etc)
Train Time Tables	station
Weather Reports	address or region
Hotel Guides	nearest station

Another type of local information is the static file. A manual analysis of static files is show in Table.2. We investigated the existence of Japanese address string for 100 files obtained randomly from the Internet. About 25% of files contains address strings. An automatic method of collecting and structuring these files are described in the section 4 of *Location-Oriented Robot-based Search*.

Table 2. Ratio of the files which contain address strings

Prefecture	City	Town	Chome[2]	Total (any address)
10.7%	13.7%	8.0%	4.0%	24.7%

2.2 Mobile Info Search

A goal of *Mobile Info Search* is to provide local information from the Internet by collecting, structuring, organizing, and filtering in a practicable form. To utilize such local information, *Mobile Info Search* employs a mediator architecture; a software agent between users and the information sources. The architecture of *Mobile Info Search* is shown in Fig. 1. Between users and information sources, *Mobile Info Search* mediates Database-type resources using *Location-Oriented Meta Search* and static files using *Location-Oriented Robot-based Search*. To use *Mobile Info Search*, only Internet connectable PDSs or PCs with Web browsers are required for users. Additionally if the user have a PHS or GPS unit, the user location is automatically obtained [2].

[2] "Chome" is a block number of the town. Japanese addressing method is Prefecture, City, Town, "Chome", "Banchi" (smaller block no.), "Go" (house no.)

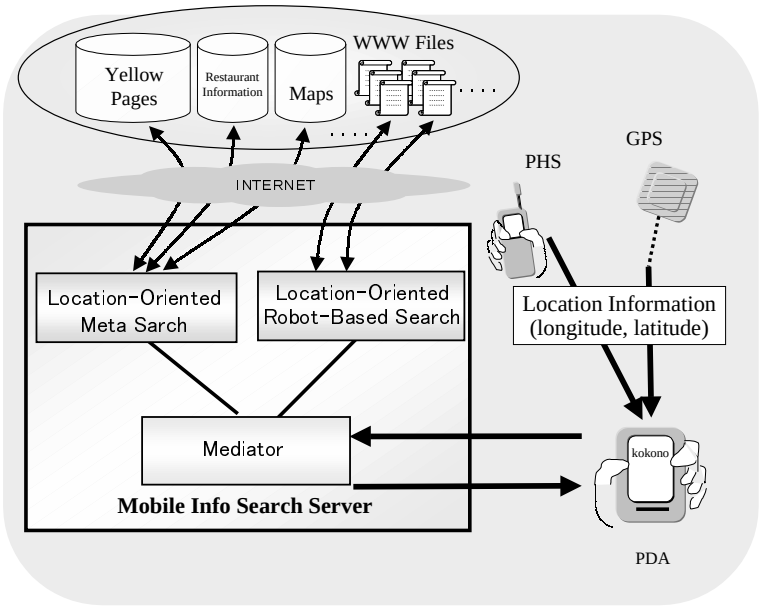

Fig. 1. An architecture of *Mobile Info Search*

3 Location-Oriented Meta Search

Location-Oriented Meta Search provides a mediation service for database-type resources. It provides a simple interface for local information services which have various search interfaces. Only to select the service such as the maps or the yellow pages, users can get information of the location easily from each server. The architecture of *Location-Oriented Meta Search* is shown in Fig. 2.

Search Controller converts the location information and picks the suitable *Wrapper* for the requested service. Both location information from the user and for the information sources are in various form. It can be address strings, longitude-latitude, postal-codes, or landmarks. The *Controller* converts the user location into the location information suitable for the target using the *Location Information Repository*. The repository we constructed on *Mobile Info Search* is the set of location information that can convert each format to others. The *Wrapper* is a software prepared manually for each services. The access method, the request form, and the usage of local information of the target is defined.

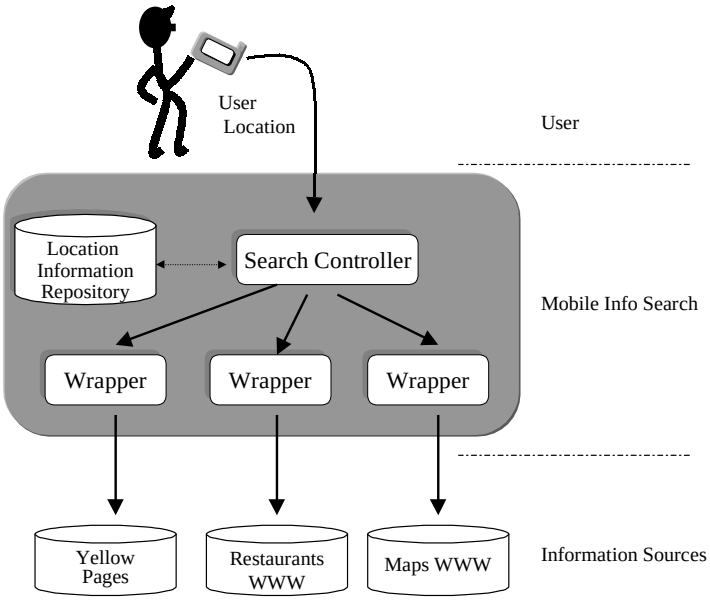

Fig. 2. An architecture of *Location-Oriented Meta Search*

Location-Oriented Meta Search performs in following way;

1. to receive user location (longitude latitude) and the name of the target service
   $ex.\ (x = 139.36.27, y = 35.25.24, target = AYellowPages)$
2. to converts the user location into the location suitable form for the target using*Location Information Repository*
   $ex.\ (x = 139.36.27, y = 35.25.24) \rightarrow address = \text{“}Tokyo, Chuo-ku, Ginza\ 4''$
3. to determine the search scope of the location for the target
   $ex.\ \text{“}Tokyo'', or\ \text{“}Tokyo, Chuo-ku'', or\ \text{“}Tokyo, Chuo-ku, Ginza'', ...?$
4. to create a query
   $ex.\ http://www.AYellowPages....co.jp/seach.cgi?address = \text{“}Tokyo, Chuo-ku, Ginza\ 4''$
5. to search

4 Location-Oriented Robot-Based Search “kokono Search”

4.1 What is kokono Search

In this section, we introduce the *Location-Oriented Robot-Based Search* called *kokono Search*. “kokono” is a Japanese word means here. *kokono Search* provides the spatial search that searches the document close to a location. Just like other

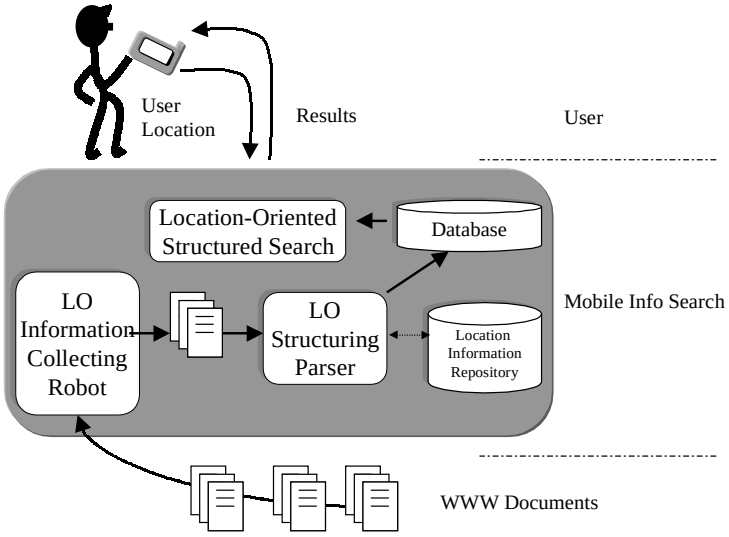

Fig. 3. An architecture *kokono Search*

search engines, *kokono Search* employs a software called "robot" that collects documents from the Internet and creates local database for collected documents. While other search engines provide a keyword-based search, *kokono Search* do a location-based spatial search. It displays documents in the order of the distance between the location of the document and the user's location. Fig. 3 illustrates the architecture of *kokono Search*. A brief flow of *kokono Search* described below.

1. to create a location oriented database
   - *Location-Oriented Information Collecting Robot* collects documents from the Internet
   - *Location-Oriented Structuring Parser* parses the obtained document to look up the location information (address strings) and store the documents with the spatial information (longitude-latitude) by consulting the *Location Information repository*
2. to perform a search
   - to search documents within the distance from the requested location
   - to display the documents in order of the distance

4.2 How to Collect Local Information

Our robot is a special one to collect local information selectively. The ability of *Location-Oriented Information Collecting Robot* is shown in Fig. 4. *Location-Oriented Robot* employs the heuristics to calculate the collection priority. The priority for the uncollected document is determined by estimating the resource is

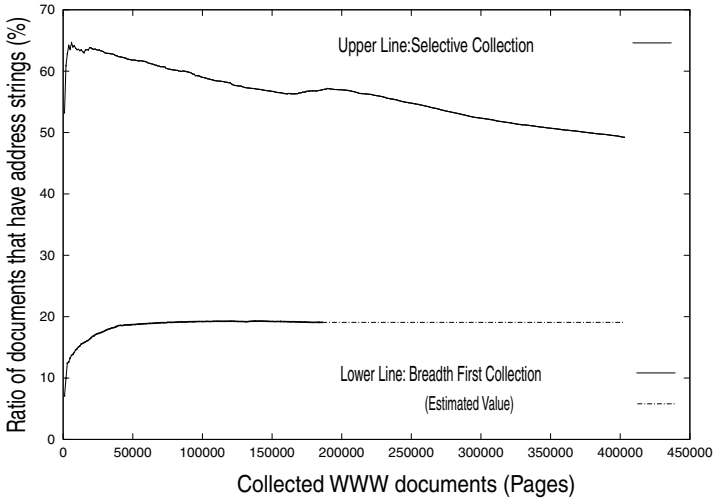

Fig. 4. Comparison between the selective information collection by *Location-Oriented robot* and the breadth first collection

the local information or not. It is estimated local information, if the link letters contain the address strings . For the starting period of collection, this robot works more effective than the normal breadth-first collection robot.

4.3 How to Structure the Local Information

To put the spatial information such as longitude-latitude to the document for the spatial search, we parse the documents to look for the location information. Location information can be address strings, station names, landmarks (ex. "Tokyo Tower"), or postal-codes. In this subsection we describe the extraction of address strings.

We extracted the address strings by following way. There are two major difficulties for address strings extraction; the writing variation and the ambiguity to the other nouns (person's name). We used several heuristics of the address strings.

1. to divide document into morphemes by the Japanese text parser
2. to compare noun phrase to the address dictionary (part of *Location Information Repository*) and regard it as an address if it satisfies the following conditions
   - any address strings without omitting the upper address, or
   - cities with address suffix (ex. Yokohama "Shi"(=City)), or
   - towns or block numbers with the city name
   - block numbers with the upper town and city name and meets the street numbers patterns (pattern ex. "1-Chome 2", "1-2", "1 2", ...)

5 Mobile Info Search: An Experimental Service

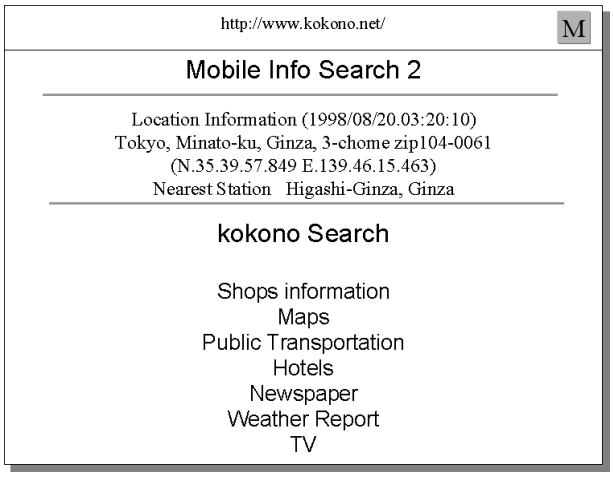

Fig. 5. Index page of *Mobile Info Search*. This page is automatically displayed by accessing *Mobile Info Search*. User's current location is displayed as the address, the longitude-latitude, and the nearest station on the top of the page. *kokono Search* and other seven categories follow. Services for the external sites are requested by selecting the services categorized into the seven menus.

Mobile Info Search is open to the public on the Internet since 1997. From the address *http://www.kokono.net/* , anyone can enjoy the *Location-Oriented Meta Search* and *kokono Search*. When user access to the *Mobile Info Search*, the index page of the location is displayed (Fig. 5). About 20 WWW services, provided by the courtesy of 8 companies are available. They include following; *kokono Search*, Yellow Pages and Restaurants guide under the Shops information, several maps, train tables, hotel guides and reservations, weather reports, and TV listings. From January to July 1999 about 500 searches a day to some services are done. Fig. 6 shows the *kokono Search* results displayed on the map.

6 Experimental Results

6.1 How They Used the Local Information? -Search Analysis-

We analyzed the user's searches from the log files recorded on the server. We used 39,718 complete searches for analysis from the log files from January to July of 1999.

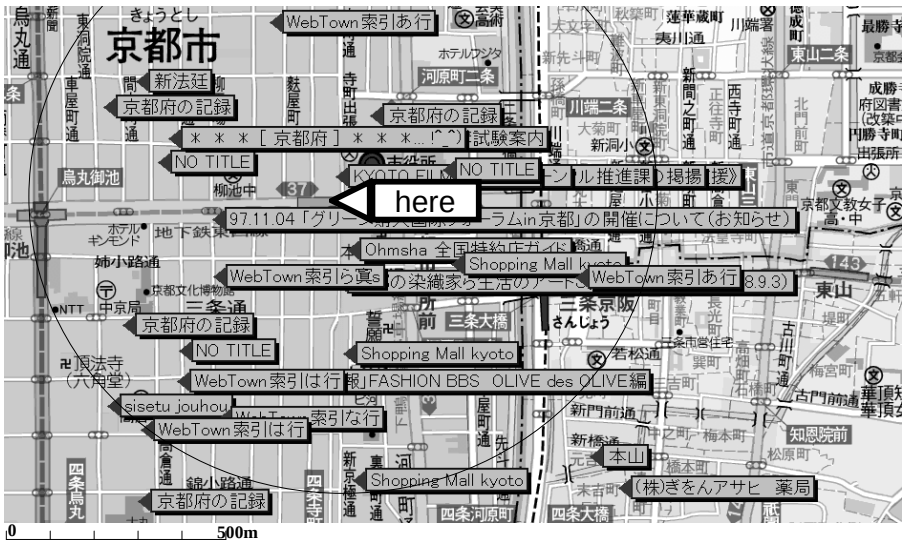

Fig. 6. *kokono Search* results at the Kyoto City Hall. Web pages are automatically located and can be retrieved geographically. All results can be displayed on the map. This is the search result of the Web pages within about 700m of Kyoto City Hall. This map is by ProAtlas (c).

6.2 Search Trends in General

Classification of all searches requested to *Mobile Info Search* is illustrates on the in Fig. 7. Map is the most requested category of *Mobile Info Search*. Note that the frequency of the request can depend on the user interface. As the map button is located at the third position of the menu and not in highlighted position, map is the basic service for the local information services. Transportation category represents the information about the rail-way station including the time table of the train. This demand reflects the Japanese living style.

6.3 Search Session Analysis

We divided searches into "sessions". We defined a session that a session is a set of searches from one person for one location in succession. If the interval of two searches was less than 3 minutes, they were regarded in succession. 39,718 searches were divided into 20,080 sessions. In about 40% of sessions, plural searches such as map and *kokono Search* are requested together and the average search of plural session is 3.44 times per session (Table. 3). Though this is a open WWW service and user may be a first-time, it is not easy to make users request plural services continuously. A sophisticated and concise way of integrating information from plural source into one screen is required.

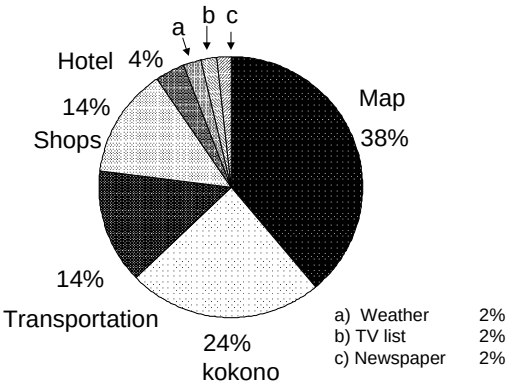

Fig. 7. Search trend of *Mobile Info Search*

Table 3. Session trend of *Mobile Info Search*

	sessions	searches/session
single search	12,047	1.00
plural searches	8,033	3.44
total	20,080	1.98

6.4 Is Digital City Created? -Data Analysis-

We analyzed about half million (479,669) documents collected for *kokono Search* by our robot and found about 40 % (210,369) of documents contained address strings (Table 4). As some documents contains more than one addresses, we collected 1,122,380 total addresses.

The collection rate for addresses is shown in Table. 5. We collected almost every cities, 30.3% of towns, and 6.8% of Chome level addresses of Japan.

Collecting some deeper addresses from the Web pages is not a easy matter. Because some Chome or deeper level addresses are used only for describing the location for personal houses, they may not appear on the Internet.

Table 4. Numbers of collected documents

	Prefecture	City	Town	Chome (Block #)	Total
Documents that have addresses (A)	210,369	151,272	68,381	32,755	[3]210,369
Addresses collected total (B)	418,488	501,992	117,066	88,926	1,122,380
Addresses per documents (B/A)	2.0	3.2	1.6	2.6	5.3

Table 5. Collection rate of documents for Japanese addresses

		Prefecture	City	Town	Chome
All addresses in Japan[4]	(A)	47	3,883	121,172	343,269
Addresses collected unique	(B)	47	3,876	36,799	23,326
Collection rate	(B/A)	100 %	99.8%	30.3%	6.8%

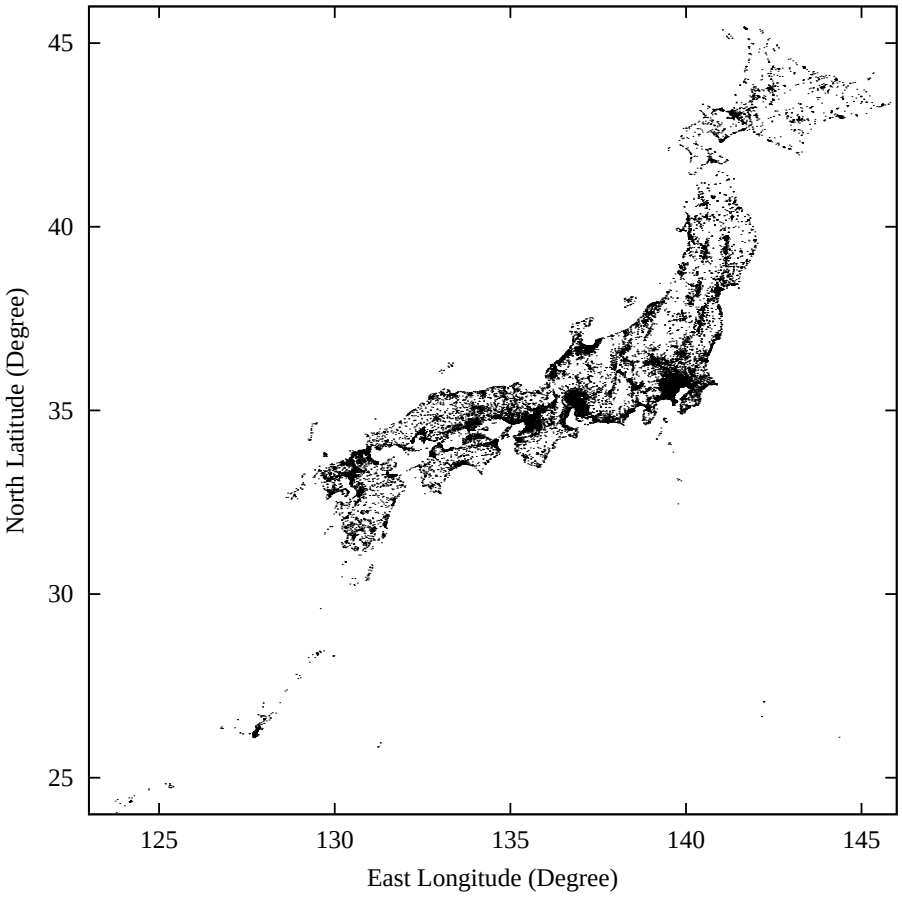

Fig. 8. A distribution map of WWW documents of Japan

Fig. 8 illustrates the distribution of the local information we collected. This map is drawn only by plotting the longitude-latitude we put to the each documents (210,369 documents). This map shows that the location described on the WWW reflects the population distribution.

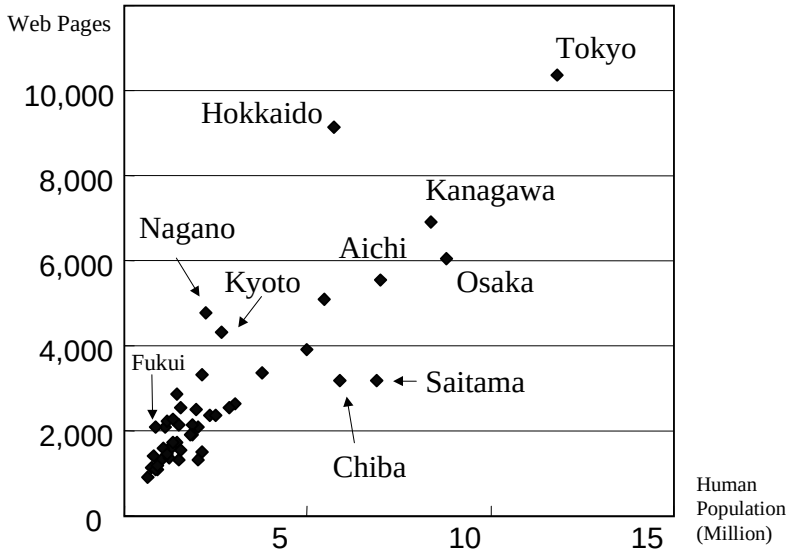

Fig. 9. Relation between human populations and Web pages

Fig. 9 shows the relationship between the populations and the number of Web pages of the prefecture. The dots represent all the 47 prefectures in Japan. We divided Web pages that have addresses (City, Town, Chome) into 47 prefectures they belong to. From the figure, there is a correlation between them. It is interesting that Hokkaido and Kyoto have many famous tourist cities. Nagano is famous for Olympic game. On the other hand, Saitama and Chiba are "bedtowns" of Tokyo and it is said most cities are common ones.

7 Related Work

We will mention our related work from the point of information search services and information mediators.

For years, the role of supplement tools for finding something in the city has been left to books or newspapers. Internet has become a useful tool for the

[3] As some document contains different level addresses together (ex. City and Town), total value is less than the sum of City, Town, and Chome

[4] from the Address Code Table published under the Ministry of Home Affairs

city since around the mid '90s. At first, it started as a searching tools / sources. Digital maps including car navigation systems and Yellow Pages services are ones of the pioneers. Our series of real-world related studies started as development of a Yellow Pages server (Internet TOWNPAGE *itp.ne.jp*) since 1995. Succeeding the growth of the real-world related contents, we have been studying information retrieval for mobile users. Intelligent Page [3] was our first effort. We intended software agents to retrieve information from the diverse resources on behalf of the user. Our next effort, Action Navigator [4] used a "recommender architecture" [5] based on the choices made by other users.

Today the activity in this area is distinguished as a "local portal sites" commercially. They collect and organize local information and provide local advertisements. Sidewalk (*sidewalk.com*), Yahoo Get Local (*local.yahoo.com*), and DigitalCity (*digitalcity.com*) are examples.

Our work is based on the work of others.

The idea of overlapping the real world and the information from the network is studied in the field of mobile computing and augmented reality[6][7].

Mediator [8] is a middle-ware service between end users and data resources. It integrates diverse data from multiple sources, reduces it to the appropriate level, and restructures the results. TSIMMIS [9][10] and I3[11] are the major mediation research projects.

There are several Internet-based research applications similar to *Mobile Info Search*. MetaCrawler [12] is a parallel WWW search service. It provides a single interface to the user, collects the results from various search engines, and organizes the results. Bargain Finder [13] and Shopbot [14] both find products by using servers and compare the prices and specification. These research efforts are close to *Mobile Info Search* in their motivations and goals; *Mobile Info Search* differs in its Location-Oriented information structuring and organizing.

In the research field of the digital city, several similar efforts are found. "Helsinki Arena 2000"[15] is the project that augments a real city to a three-dimensional virtual one. "Digital City Kyoto"[16] also tries to reproduce the real city and to create the a social information infrastructure on it. Both projects take the real and virtual world to be tightly linked and the virtual one to be a infrastructure of the city. We believe *Mobile Info Search* and these projects work together to realize the digital city.

8 Conclusion

We introduced Mobile Info Search. Location-Oriented Information Integration is the key word of our project. To support the interaction between the human and the city, the real-world, we are studying the method that integrates useful local information existing much on the Internet in a Location-Oriented way. We described out current two methods, *Location-Oriented Meta Search* and *Location-Oriented Robot-Based Search*. And also described about the experimental results from the test service on the *http://www.kokono.net/*.

Our approach for the Digital City effort is from the side of information retrieval and search. We can feel the city by searching local information on the net. Such information will be growing and may change their forms and roles. We will follow it by studying the technique for processing local-information sources and learn the city by looking the real-world and information together.

References

1. Takahashi, K., Miura, N., Sakamoto, H., Shima, K.: Location Oriented Information Integration. In Proc. Japan World Wide Web Conference '97, http://www.iaj.or.jp/w3conf-japan/97/ (1997)
2. Takahashi, K., Miura, N., Yokoji, S., Shima, K.: Mobile Info Search: Information Integration for Location-Aware Computing. In Proc. IPSJ SIGMBL http://www.kokono.net/~ takahasi/articles.html (1998)
3. Takahashi,K., Nishibe,Y., Morihara,I., Hattori,F.: Intelligent Pages: Collecting Shop and Service Information with Software Agents. Applied Artificial Intelligence Intl. J., Vol. 11, No. 6, Taylor & Francis (1997) 489-499
4. Ohtsubo, R., Nishibe, Y., Takahashi,K.,Morihara,I.: Action Navigator: an Information Service Based on Agent-Communication for Supporting Decision-Making. In Proc. PAAM98 (1998) 629-630
5. Resnick,P., Varian,H.R.: Recommender Systems. CACM, Vol.40, No.3 (1997) 56-58
6. Wellner. P., Mackay, W.E., Gold, R.: Computer-Augmented Environments: Back to the Real World. CACM Vol.36, No.7 (1993) 24-26
7. Nagao, K., Rekimoto, J.: Agent Augmented Reality: A Software Agent Meets the Real World. In Proc. the Second Intl. Conf. on Multi-Agent Systems (ICMAS-96) (1996) 228-235
8. Wiederhold, G.: Interoperation, Mediation, and Ontologies. FGCS '94 Workshop on Heterogeneous Cooperative Knowledge-Base, (1994) 33-48
9. Hammer, J., et al.: Information Translation, Mediation, and Mosaic-Based Browsing in the TSIMMIS System. ACM SIGMOD Intl. Conf. Management of Data (1995)
10. TSIMMIS: The Stanford-IBM Manager of Multiple Information Sources (TSIMMIS). http://www-db.stanford.edu/tsimmis/ (1998)
11. Gunning,D.: Intelligent Integration of Information Technology (I3). http://maco.dc.isx.com/iso/battle/i3.html (1997)
12. Selberg,E., Etzioni,O.: The MetaCrawler Architecture for Resource Aggregation on the Web. IEEE Expert, Volume 12 No. 1 (1997) 8-14
13. Krulwich, B.: Bargain finder agent prototype. Technical report. Anderson Consulting. http://bf.cstar.ac.com/bf/ (1995)
14. Doorenbos, R., Etzioni, O., and Weld: A Scalable Comparison-Shopping Agent for the World-Wide Web. Autonomous Agents '97 (1997)
15. Linturi, R., Koivunen, M. R., Sulkanen, J.: Helsinki Arena 2000: Augmenting a Real City to a Virtual One. Lecture Notes in Computer Science (in this volume), Springer-Verlag (2000)
16. Ishida, T., et al.: Digital City Kyoto: Towards A Social Information Infrastructure. Cooperative Information Agents III, Lecture Notes in Artificial Intelligence, Vol. 1652, Springer-Verlag (1999) 23-35