MyMemex: A Web Service-Based Personal Memex System

Youngkun Min and Bogju Lee
Department of Computer Science and Engineering
Dankook University
Yongin City, Republic of Korea
{minyk, blee}@dankook.ac.kr

Chansu Yu
Department of Electrical and Computer Engineering
Cleveland State University
Cleveland, Ohio
c.yu91@csuohio.edu

Keywords*-memex; personal memex; life log; MyLifeBits, web service; ontology; memex event; episode classification*

I. THE MYMEMEX SYSTEM

MyMemex server consists of a web server, a data collection agent, a file handler, and a database. The data collection agent makes connections to the company web services and stores the collected "web data" (phone logs, credit card usage logs, e-mails, and so on) to the memex database. The web server enables the users to view the collected data and get the results for the queries. The users can also upload the "file data" such as image, video, and audio files locally stored in user's personal computer. Before saving the memex data, the web server converts the various types of data into the standard 4W1H form. The "memex event ontology" is used in the conversion.

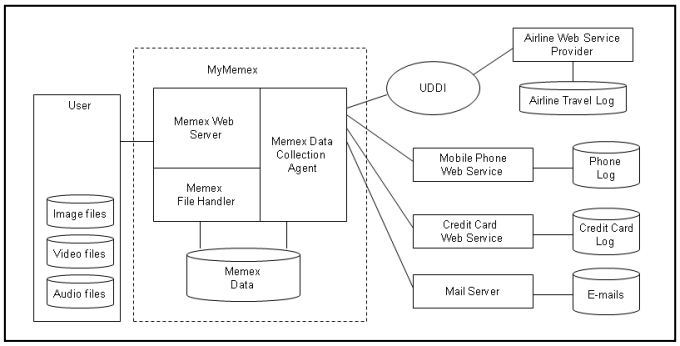

Figure 1. The archirecture of MyMemex

Related events tend to happen consecutively. A group of related events is called an "episode". Given a series of memex events, MyMemex employs a heuristic algorithm to solve the episode classification problem.

II. USER INTERFACE

MyMemex provides two types of data view – "4W1H view" and the "story view". 4W1H (when, where, who, what, and how) view is shown in Fig. 2. Memex data are shown in 4W1H form chronologically. Users can navigate through old and new data and simple keyword-based search is available.

The story view provides the user with a view of diary or article styles. Each memex datum is converted to one sentence whereas "how" becomes another separate sentence. Each episode forms one paragraph.

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2008-314-D00361).

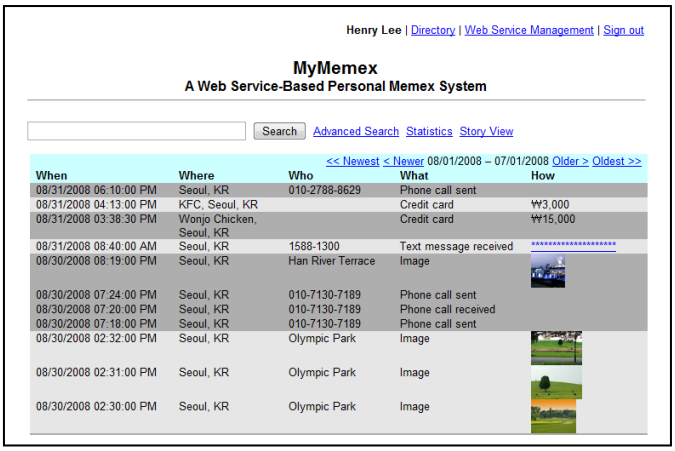

Figure 2. 4W1H view of the memex data

III. EPISODE CLASSIFICATION

Our episode classification algorithm is based on the following heuristics: (1) the memex data within a small time window (i.e., their when's are very close) are in one episode; (2) if the data is within a bit larger time window and have the same "who" or "where", they are in one episode. The data that their episode ID's are assigned to by the algorithm are saved into the database. Then the heuristics is applied again among the remaining data where episode ID's are not given.

Table 1 shows the number of total data, the number of correctly classified data. The accuracy indicates that the heuristics is performing relatively well.

TABLE I. EPISODE CLASSIFICATION RESULT

The number of memex data	*The number of correctly classified*	*Accuracy*
494	445	0.90

REFERENCES

[1] V. Bush, "As we may think," The Atlantic Monthly, vol. 176, pp. 101-108, July 1945.

[2] J. Gemmell, G. Bell, and R. Lueder, "MyLifeBits: a personal database for everything", Communications of the ACM, vol. 49, pp. 88-95, Jan. 2006.

[3] K. Aizawa, S. Kawasaki, D. Tancharoen, T. Yamasaki, "Efficient retirieval of life log based on context and content", ACM Multimedia Workshop on Continuous Archival of Personal Experience 2004, pp. 22-31, Oct.15, 2004, NY.