

PaperWeb: Paper-triggered Web Interactions

Sriganesh Madhvanath, Geetha Manjunath, Suryaprakash Kompalli, Serene Banerjee, Sitaram Ramachandrula, Srinivasu Godavari
Hewlett-Packard Laboratories
24 Salarpuria Arena, Hosur Main Road
Aduodi, Bangalore 560030, India
+91 80 3382 9169
srig@hp.com

ABSTRACT

While mobile phones have penetrated deep into tier 2 and 3 cities in India and similar emerging economies, adoption of mobile web content and web services is likely to require the creation of large numbers of relevant applications and services with usable interfaces and interaction paradigms. This paper describes *PaperWeb*, our effort to enable mobile phone users to use the web for day to day transactions such as paying bills, buying tickets, or fixing appointments, using familiar objects such as paper artifacts. We discuss (i) the creation of useful PaperWeb interactions – without programming – by moderately tech-savvy users, and (ii) the use of these interactions by tech-naive users, and briefly describe the underlying technology. We conclude the paper with a discussion of current status and next steps.

Categories and Subject Descriptors

H5.2. Information interfaces and presentation: User Interfaces.

General Terms

Design, Human Factors.

Keywords

Web tasks, mobile cameras, document image processing

1. INTRODUCTION

The web is one of the greatest inventions of the 21st century and its end-user value just in terms of productivity – tasks such as booking tickets, looking up information – is being reaped by millions of people around the world who can use a web browser. However despite the rapid adoption of mobile phones in emerging markets (approximately 15 million subscriptions being added each month in India alone), large numbers of individuals and small businesses in these geographies remain unable to harness the power of the web, because using a browser is often too complex a task, requiring considerable understanding of web page addressing and navigation and other concepts that we take for granted.

This was validated in a recent ethnographic user study of middle-income households in the Tier 2 town of Bhopal, where we discovered that despite a number of services moving online and these families being able to afford computers and smartphones, routine tasks such as paying bills or booking tickets still require standing in lines and considerable expenditure of time and money.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
DEV '13, January 11-12, 2013 Bangalore India
Copyright © 2013 ACM 978-1-4503-1856-3/13/01... \$15.00

In this paper we describe *PaperWeb*, a system that allows users to take pictures using their mobile phone of a *paper artifact* and perform an associated *web based task*. For example clicking a picture of a utility bill enables the web task of paying the bill for that particular month. We discuss the authoring of such experiences by moderately net-savvy users using simple click and browse actions, which we believe can lead to a long tail of such experiences. PaperWeb combines document image processing technologies with web task automation to enable access to web services using a familiar interface.

There have been a number of attempts over the years at relating physical objects to digital information, from Digital Desk [2] to Sixth Sense [3]. There are also specific mobile applications available today for reading barcodes from books and other products and fetching online reviews and product information. Google Goggles [4] enables generalized web search from popular images, but does not address triggering specific tasks from personally relevant documents. We believe that our work is the first to focus on the personal productivity needs of non tech-savvy users in emerging economies, and to propose a general authoring-based solution for enabling a wide variety of web tasks based on paper (and other image) triggers.

2. THE PAPERWEB SYSTEM

The PaperWeb system enables relevant online experiences to be authored once by users with the ability to use a web browser and a PC (but no programming skills), and used by non-tech savvy users using camera-enabled mobile phones (or PCs with a connected camera or scanner) for document capture.

Authoring: Authoring of PaperWeb interactions is performed using a desktop PC based authoring tool and involves three steps. The *first* step is the *authoring of task-specific web widgets called tasklets* using a “programming-by-doing” metaphor [1]. In simple terms, it enables recording the sequence of web browsing actions required to carry out the online task (e.g. paying the utility bill) and automatically packages it into a web widget (e.g. *PayBill*) that executes in the cloud when triggered. Any user inputs required by the task (e.g. *bill date*, *amount*) become *parameters* and are prompted for when the web task is subsequently executed. These web tasks may be tagged as *private* (for the author’s use only) or *public* (may be used by others), and are published to an online *web task repository*.

The *second* step of authoring is *document template authoring*. Using a single scanned sample of the paper artifact that might be a natural trigger for the web task (e.g. utility bill), all *fields of interest* (e.g. *acct number*, *bill date*, *bill amount*) are marked to create a *document template*, which is then published to a cloud-based *document template repository*. The *final* step of authoring involves associating a web task (private or public) from the web task repository with a document template from the document



Figure 1: Usage of TaskGenie client application on mobile phone: (i) User takes a picture of her railway ticket, (ii) document is automatically identified as a railway ticket, and previously authored and mapped web tasks (booking status inquiry, cancel ticket, etc) are retrieved, (iii) on selecting the booking status task, the required web task parameters are automatically extracted from the document, (iv) the web task executes and booking status from the railways website is displayed.

template repository, and (optionally) specifying the mapping between web task parameters (e.g. *amount* parameter of *PayBill*) and fields of interest in the template (e.g. *bill amount* field on the utility bill). A web task may be mapped to more than one template (and potentially by different users), and vice versa.

Usage: As shown in Figure 1, the end user (who in general is different from the author) uses a client application called *TaskGenie* on his mobile device to access and execute web tasks. Using this client, the user can search the task repository for tasks of interest using keywords. Alternatively, she can capture a physical document (e.g. the utility bill that arrived in the mail) using the mobile camera and retrieve tasks associated with that document. Behind the scenes, the image is transferred to the cloud, and the input document identified from among all document templates registered with the repository. The user then selects one of the retrieved web tasks to execute. This causes the web task parameter values to be obtained automatically from the corresponding fields on the document, obviating the need for manual entry of these values. The user can correct the automatically extracted parameter values. Thereafter the web task executes in the cloud and returns the results to the user. For the end user, the familiar paper bill thus enables her to use the web to pay the bill with a single click.

3. COMPONENTS AND TECHNOLOGIES

The key components of the PaperWeb system are: (i) the desktop authoring tool that supports creation and association of web tasks and document templates, (ii) the cloud-based web task repository (which also stores the document associations), execution engine and task portal, (iv) the cloud-based document template repository and image processing services for document identification, field extraction and OCR (v) TaskGenie client applications for different mobile and desktop platforms.

The PaperWeb system integrates several image processing and web modeling algorithms. An algorithm based on compressed sensing is used for identifying the input document given the document templates provided during authoring. Extracting web task parameters automatically from the paper document involves aligning the input image with the identified document template, extracting field sub-images, and interpreting them using an OCR engine. These steps have been implemented using algorithms designed to be robust in the presence of variations in image illumination, skew and perspective distortion, cropping and blur.

We have also made tasklets more robust with respect to web site changes through multi-way modeling of websites, detection of semantic errors and proactive/reactive ways of identifying and characterizing web changes. The system also integrates support for "live" (long-running) tasklets with a "semantics-based" cache for handling disruptions in connectivity.

4. DISCUSSION AND NEXT STEPS

We have prototyped the PaperWeb system using a PC-based authoring tool, TaskGenie clients for multiple mobile and desktop environments including Windows, Android, Windows Mobile and WebOS, and cloud-based web task and document template repositories hosted on Amazon Web Services. We have authored multiple web tasks such as train ticket booking, mobile bill payment and loyalty points balance checking and mapped them to the corresponding ticket, bill and loyalty card templates. We have tested our prototype using multiple imaging devices such as mobile cameras, desktop scanners and document cameras.

In initial tests, document identification from poor quality mobile camera images has proved to be 90% accurate for a repository of ten document templates. Document field extraction is 96% accurate for common image distortions. OCR accuracy is greatly dependent on the availability of autofocus in the camera and a flash for low light conditions, and the presence of handwritten content and patterned backgrounds (e.g. on a credit card).

The overall system is adequate for a small set of personal interactions; however in order to support sharing of document templates and authored PaperWeb interactions with other users, it will be important to build in error recovery especially for poorly captured images, and scale both robustness and performance of the algorithms to deal with hundreds of document templates. Authoring techniques and algorithms for dealing with documents that do not follow a fixed template (e.g. restaurant bills) need to be incorporated. Also of interest are use cases relevant to small and medium enterprises, and those involving image triggers other than documents (e.g. movie posters, street signs). Key next steps include field testing with end users to understand usage issues once deployed in the real world, and exploring service-oriented business models for the rollout of such services.

5. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Vimal Sharma and Nidhi Mathur who led the user research and user experience design for PaperWeb, and S. Thara, A. Jayakumar, and Ranjeetha Venkatesh who led the software prototyping and system integration effort.

6. REFERENCES

- [1] Manjunath, G., Guntupalli, S., Thara S, Vinay Kumar K., Bosamiya, H., Ragu Raman G. 2009. Creating Personal Mobile Widgets without Programming, In *Proc. WWW 2009*.
- [2] Wellner, P. 1993. Interacting with Paper on the DigitalDesk. *Communications of the ACM*, 36, 7 (Jul. 1993), 87-96.
- [3] Mistry, P. and Maes, P. 2009. SixthSense: A Wearable Gestural Interface. In *Proc. ACM SIGGRAPH ASIA 2009*.
- [4] Google Goggles. <http://www.google.com/mobile/goggles>.