



Prof. Dinkar Sitaram

Prof., Department of Computer Science and Engineering, PESIT
Head, Center for Cloud Computing and Big Data

MULTI-MODAL PHONE INTERFACES



Overview

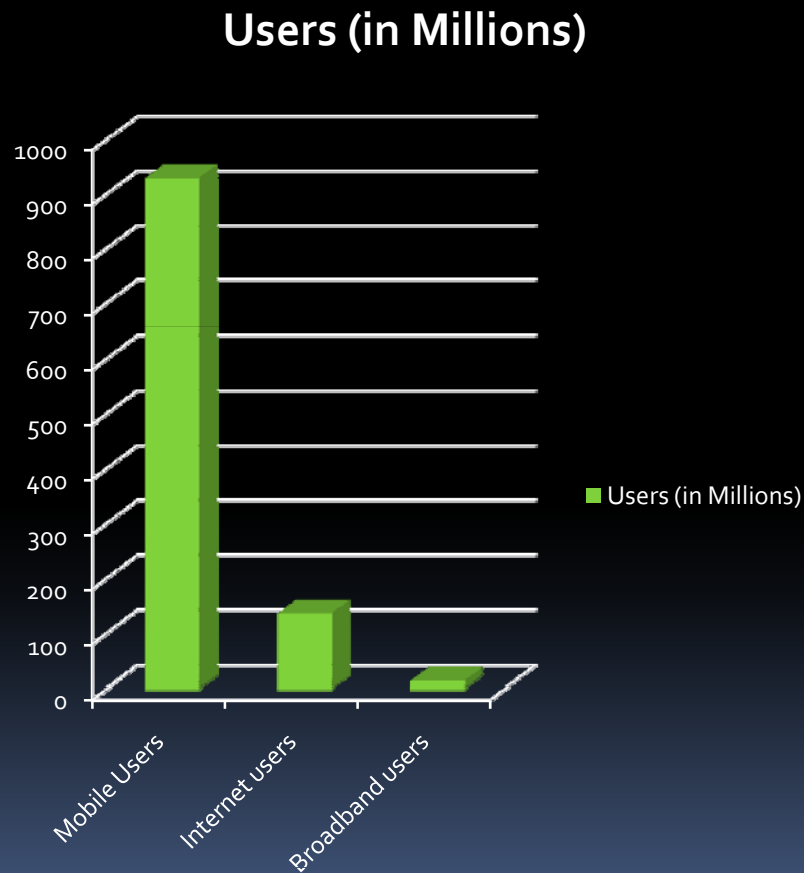
- Introduction
 - Survey of work
- 



Introduction



Phone vs Internet in India



- Broadband: dedicated line to house
- Person who does not have Internet access cannot
 - Google (e.g., health info)
 - Access web information
 - Online booking
 - Government info

How can Internet Usage be Increased? – 1/2

- Smartphones are not the answer
 - Expensive
 - Large
 - Users not computer-literate
 - Semi-literate users
- Problem: Internet is designed for use by PCs
 - Text-based interfaces cumbersome for mobile device
 - Explains popularity of Siri

How can Internet Usage be Increased? – 2/2

- Preferred interface for phones
 - Speech
 - Simple button clicks
 - 'Press OK to accept'
- VoiceXML isn't a solution
 - Creates 'speech ghetto'
 - Disconnected from Internet

MIT Technology Review

NEWS & ANALYSIS · FEATURES · VIEWS · MULTIMEDIA · DISCUSSIONS · TOPICS · POPULAR: SOLAR DOWNTURN

Is this site looking too big? The site automatically scales based on the font size



Forum and Exhibition

VIEW

4 COMMENTS



Tom Simonite
February 8, 2013

Siri, Google Now, and the End of Apps

If Siri gets more powerful people will have little use for many apps.



User Interface Study

- Experimental study
 - Textual interfaces unusable by non-literate users
 - Error-prone for novice literate users
 - Speech effective for 70% of users
 - Graphical interfaces helpful for all users
- “Designing mobile interfaces for novice and low-literacy users”
 - Indrani Medhi (Microsoft Research, India) Somani Patnaik (MIT) et. al.,
 - ACM Transactions on Computer-Human Interaction (TOCHI) Volume 18 Issue 1, April 2011

Technology Challenges

- Speaker-independent continuous speech recognition
- Simplified user interaction
 - Multiple links from a single Web page
 - Present only the most important
 - Link selection via speech
- Simplified presentation of Web page
 - Web page intended to be displayed visually
 - Rich text and data
 - Extract important features
- All the above require sophisticated semantic analysis

Speech Technologies for mobile usage

- Support for accent agnostic recognition for Indian languages and dialects
- Support for noisy and suboptimal acoustic conditions
- Mixed language interactions
 - eg. English + Kannada
- Language modelling to accommodate grammatical inaccuracies, incompleteness



Survey of Work



Novice User training

- Joint effort by CMU and Lahore Institute of Management Sciences
- Game
 - User dials number
 - Instructions to
 - Record message
 - Playback
 - Forward to friend

The screenshot shows the MIT Technology Review website. At the top, there is a navigation bar with the MIT Technology Review logo and a menu with items: NEWS & ANALYSIS, FEATURES, VIEWS, MULTIMEDIA, DISCUSSIONS, TOPICS, and POPULAR: SOLAR DOWNTURN, SAPPHIRE REPLACES GLASS. Below the navigation bar is a yellow banner with the text: "Is this site looking too big? The site automatically scales based on the font size set in your browser." Below the banner is a white box with the MIT Technology Review logo and the text: "Mobile Summit | June 10-11, 2013 San Francisco, CA REGISTER TODAY". Below the white box is a section titled "COMMUNICATIONS NEWS" with the article title "Viral Phone Game Helps Illiterate Pakistanis Find Job Listings". The article text reads: "A viral phone game in Pakistan trains people to use their keypad—and gives them the skills they need to hunt for a job." On the left side of the article, there is a vertical sidebar with icons for social media and other functions.

Voice-Based Social Media for Developing Regions

- Stanford Department of Computer Science
- Social media for rural communities
- Challenges
 - Limited capability speech recognition for Indian languages
 - Designing for illiterate users
 - Search and filtering of audio content

Hearsay

- Multi-modal non-visual web browser
 - Stony Brook University
 - Envisioned as next-gen assistive technology for web access
 - Supports multiple output (audio, visual, braille) and input (speech, keyboard, touch, phone keypad)
 - Incorporates information retrieval, machine learning and content analysis, statistical language detection
 - Support for personalized speech-enabled macros for automatic repetition of tasks
 - Multi-modal interactions for quicker and comfortable interfacing

Site on Mobile

Site-On-Mobile

Unveiling a novel technology that enables the value of the Internet to reach the masses



End-User Programming

Packaging Frequent Web Interactions



System
generated
widgets



Desktop
Widgets



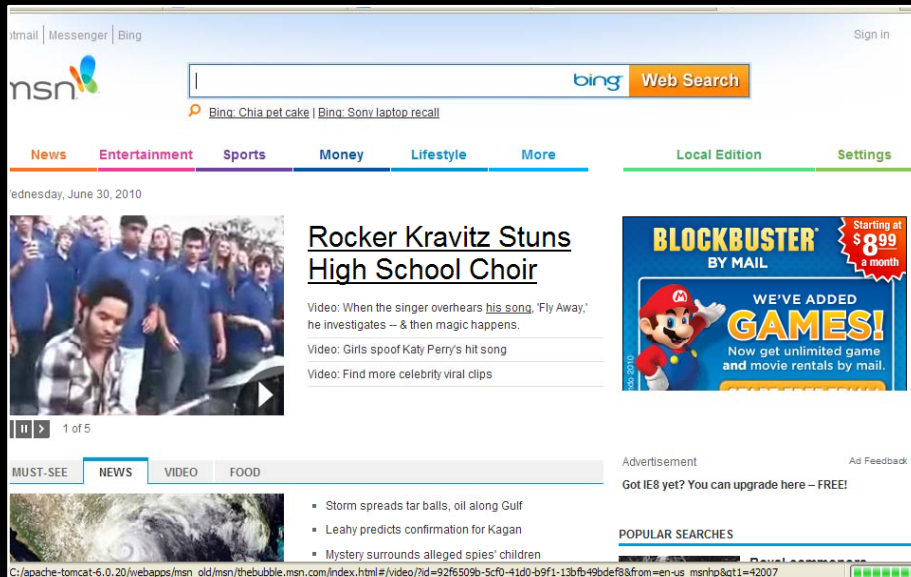
Mobile
widgets, SMS, Voice

Task Specification

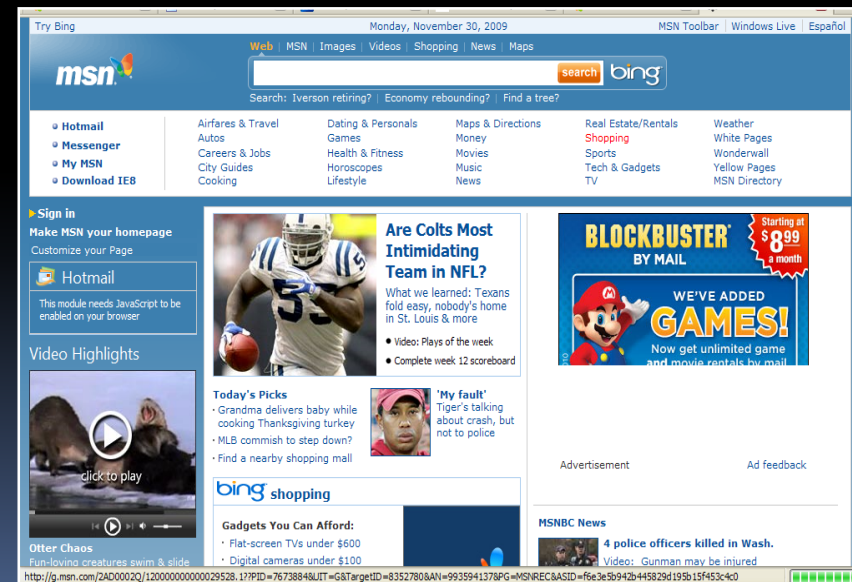
Observe/Record User Actions

Programming-By-
Demonstration



Web Site Changes Examples



Different versions of
<http://www.msn.com>
over one year





Types of Web Changes

	<h2>HP Pavilion Home Desktop PCs</h2>
 starting from <u>Rs. 30,990</u> Taxes & levies extra	<p>HP Pavilion Slimline a6600 Desktop PC series</p> <ul style="list-style-type: none">The perfect home PC for photos, music, videos and surfing together

ORIGINAL PAGE

<h2>HP Pavilion Home Desktop PCs</h2>	
<p>HP Pavilion Slimline a6600 Desktop PC series New!</p> <ul style="list-style-type: none">The perfect home PC for photos, music, videos and surfing together	 starting from <u>Rs. 30,990</u> Taxes & levies extra

LAYOUT CHANGE

	<h2>HP Pavilion Home Desktop PCs</h2>
 starting from <u>Rs. 30,990 (Festival discount 10% in this series)</u> Taxes & levies extra	<p>HP Pavilion Slimline a6600 Desktop PC series</p> <ul style="list-style-type: none">The perfect home PC for photos, music, videos and surfing together

CONTENT CHANGE

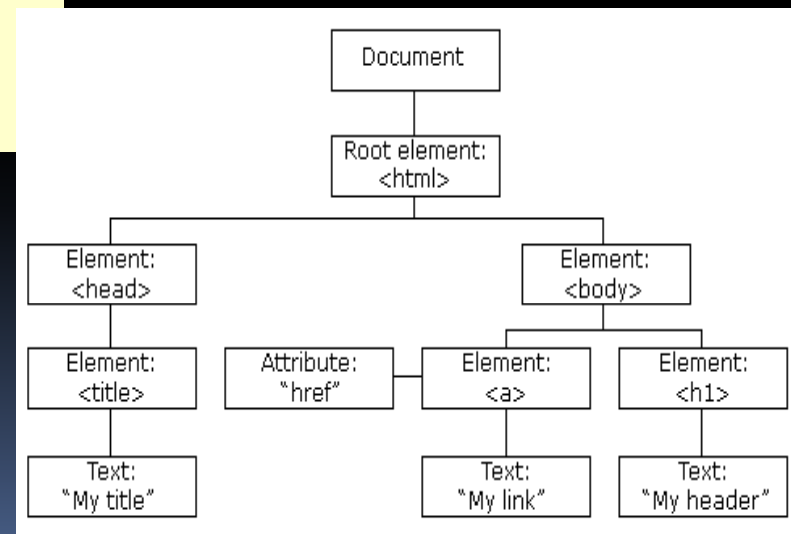
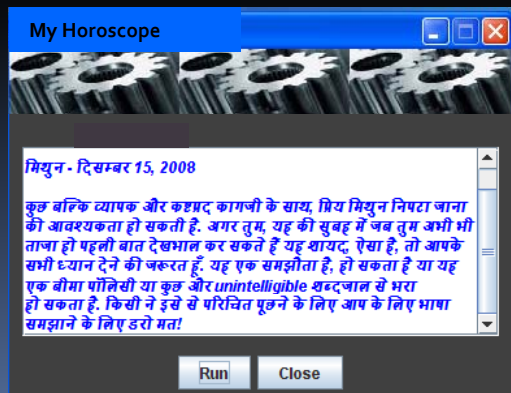
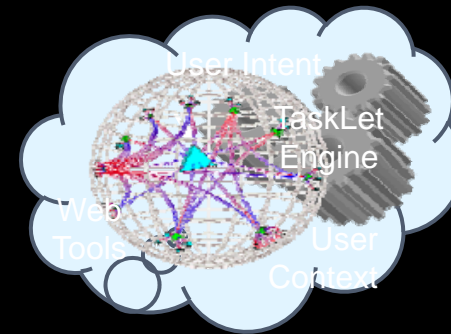
Others: Attribute Change, Semantic Change, Input Error, Surprise Captcha,...

TaskLet Execution

Programming-the-Web despite the website changes



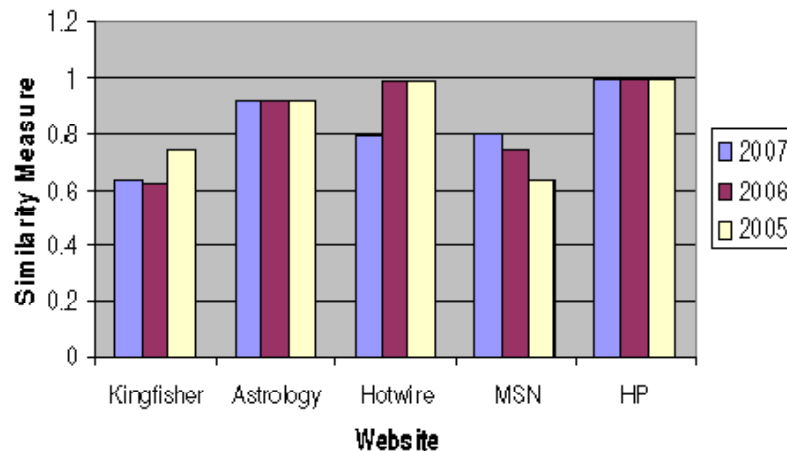
- Dynamic TaskLet Engine
 - A Virtual Machine with Web Opcodes
 - Change Handling for every interaction type
- Multi-way Modeling of Web Sites
 - Structural Modeling (DOM Structure)
 - Semantic Modeling (Spreading Activation Networks)
 - Modeling Change (Dynamic Network Analysis)
- Post Processing
 - Content Transformation using existing web services



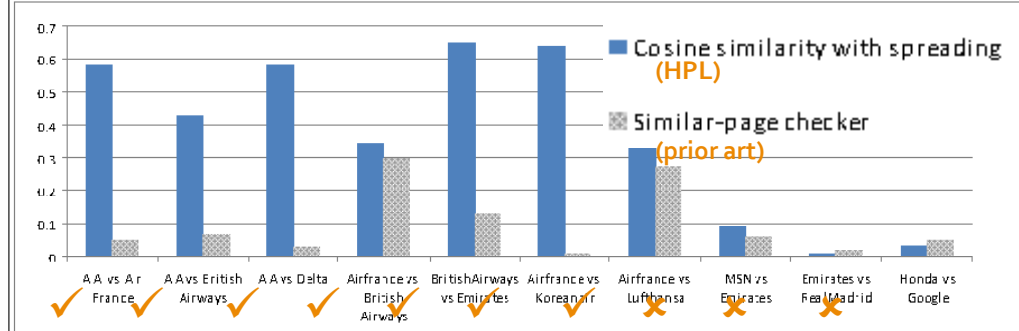
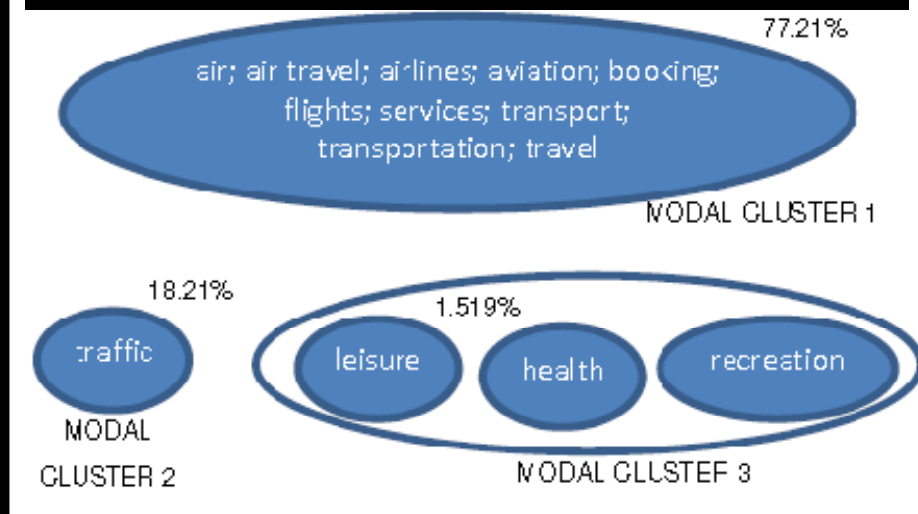
TaskLets on Similar Web Sites

- Stretching the extent of support for changed web sites
- Homogeneous vs Heterogeneous websites (using modal clustering)
- Defining a Similarity Measure on changing websites

Changing Versions of Websites



Modal Clusters to determine Heterogeneity of a website



Comparing absolute values of the similarity measure

Siteonmobile Web site

www.siteonmobile.com

The screenshot shows the website interface within a Microsoft Internet Explorer browser window. The browser's address bar displays <http://www.siteonmobile.com/>. The website header includes the SiteonMobile logo (tagged as LABS™) and a "Member Sign in" link. The main content area is titled "A technology to enable SMS and Voice services for your website, in three easy steps..." and features a three-step process:

- 1.** You select your most accessed web content
- 2.** Our technology helps you Package your web interactions (labeled "Free!")
- 3.** Your web content/service is now accessible through any mobile & landline (labeled "SMS Easy" and "Voice enabled")

An orange "Invitation Only!" button is positioned to the right of the third step. Below this process is a navigation menu with four tabs: "What is it?", "How do we do this?", "Why you need it?", and "Why us?". A carousel below the menu displays three featured items:

- Compressing Web:** Compressed web interactions to reach-out to all users (with a pill icon).
- In news...:** What media is talking about us... (with a newspaper icon).
- Sharing our story...:** Please share your comments on *Our Blog...* (with a blog icon).

The footer contains the following text: [Home](#) [About us](#) [Terms of use](#) [Privacy policy](#) [Contact us](#) © 2009 Hewlett-Packard Development Company, LP.