E-Business

Alex Sverdlov

`alex@theparticle.com`

1 Introduction

Until mid-1990s, the general public did not know much about the Internet. With the introduction of a web-browser, as well as broad advertising campaigns by Internet Service Providers (such as AOL, etc.) and useful services such as email, people started signing onto the internet. With more eyeballs came the motivation to build content. More content meant more reasons for eyeballs to sign up.

The speed of adoption was greatly aided by HTML: the language behind websites. It defines the structure of a document (web-page), and allows browsers to display it mostly how the designer indented. HTML is very easy to learn and is very forgiving of mistakes—it was widely used (and abused) to build everything.

1.1 HTTP

The HTTP protocol is a request-response protocol: where client requests a resource, and the server responds with either a resource of a code indicating an error. In many scenarios, the resources are web-pages (HTML documents).

1.2 Cookies

The HTTP protocol is stateless, which means that the server does not maintain state from request to request. Cookies were added to enable the passing of (often short) strings from client to server, with the sole purpose of maintaining a state from request to request.

Cookies often contain the session identifier, which the server stores in a database along with other session information (who is logged in, etc.). The web-browser just returns the cookie that it was given.

1.3 Security

The base HTTP protocol has no security: everything is clear text (in fact, it is possible to interact with the web-server manually via telnet).

Security is required to manage secure sessions and payment processing. A lower level protocol, SSL (secure socket layer) is often used underneath HTTP to enable encrypted connections.

1.4 Directories

With the amount of content on the internet growing exponentially, it became increasingly important to be able to find needed content. Two strategies for data organization came up: directories and search engines.

Directories are semi-manual way of organizing the information by topic. To find things, a user would traverse a tree of links from broad topics to more specific topics until the information is found. One of the first successful examples of that is Yahoo!.

1.5 Search Engines

A piece of software called a spider crawls the web—downloads a page, and queues up all links on that page for subsequent downloads. Essentially recursively downloading the entire internet by following the links.

Another program called an indexer runs through the downloads and creates an index of words to documents.

A user query that consists of several words can quickly be satisfied, essentially doing a merge join on the document lists of all looked up words.

While the above works to find documents that contain all required words, it does nothing regarding document relevancy. With lots of content, the problem is not to find documents that have words, but to find documents that are useful. There is a need to rank the search results—to present the most relevant documents first.

1.5.1 TFIDF

One of the first attempts to rank search results was to use *term frequency inverse document frequency*. Essentially a measure of importance: how often does a term show up in a document over how many documents it shows up in. So a common term (such as "the", "of", "is", "to", etc) that shows up in all documents will have a very low weight, while rare terms (such as "tfidf") that only show up in handful of documents will have a higher weight.

1.5.2 Page Rank

There were many attempts to rank pages by the web-structure. The first few attempts were counting links: in-bound links and out-bound links. The documents with many out-bound links may be labeled as hubs (they take web-surfers to other sites). Documents with many in-bound links would we weighted higher because they are apparently popular.

This setup can be gamed. A web-site designer can create lots of pages and have them point to a single page that they want to be ranked high.

Larry Page & Sergey Brin came up with rank that depends not only on the number of in-bound links, but also on the quality of those in-bound links. In-bound links of high rank give a page high-rank. Also, out-bound links share the rank: so a high ranked page cannot pass on its high rank to thousands of links. They called it 'Page Rank' (after Larry Page) and founded Google.

1.6   History: Java, Javascript, Java Applets

Around 1995, Sun came out with Java programming language. Java is object oriented language that gets compiled to byte-code, which runs on a virtual machine. This enabled programs written in this language to run on any platform that runs the virtual machine.

Javascript is an unrelated language, which enabled scripts to be included within web-pages, that the web-browser would execute.

Java Applets showed the world that rich-web-enabled applications are a possibility— maybe not a viable approach (because of speed and security issues), but it was feasible.

1.7   CGI

Whenever a web-server handles a request for resource, if the resource is a computer program, then the web-server may execute it—the way the program gets parameters and produces output that is subsequently displayed by the web-browser is standardized by CGI (common gateway interface). Any program can be made to run as CGI.

Today it is relatively common to have PHP or Java handle web-requests (through CGI).

1.8   AJAX

Javascript eventually eventually evolved into a full blown programming language, with lots of functional features and the capability to do things that were beyond its intended purpose.

One of the big things is: AJAX. This is essentially a function XMLHttpRequest, which creates a callback—allowing web-pages to communicate with the server and to change state whenever the server responds.

Much of the existing interactive web-pages are built using ajax.

1.9   Portals

Many companies created "portals" to consolidate all their services under the same entry-point. User logs in, and can do whatever the company offers using a common user interface.

1.10   Broadband

The Internet had a few speed revolutions—when consumers were first exposed to the web, via AOL or similar, often the connection was via the phone line, at 9600bps. It quickly moved to 14.4kbps then to 28.8, and topped out at 56kbps.

From there, DSL and Cable (and subsequently fiber) service enabled much faster connectivity—bringing video streaming capability to the web.

1.11 Mobile

The latest 'web' revolution is in mobile. Until about 2007, very few folks accessed the internet through their phones, and mobile apps were virtually non-existent.

Today, a large proportion of the world accesses the internet through their phones. There are millions of mobiles apps—created by an industry that did not previously exist.

2 Online business

Internet enables far wider reach than a physical store. So the first obvious business was for physical businesses to sell things over the internet.

Because the barrier to entry is so low (no need to rent a physical location), lots of online-only businesses were started.

Many businesses catered to a niche customer base—things that were hard to find in physical stores could be ordered online. Amazon was one such store: physical stores often don't have enough space to carry every book ever published, but it is often relatively easy to order online.

It is important to note that the "business" has not really changed. Everyone is still trying to sell something—now online.

3 Advertising

For all the businesses trying to sell anything online, they need customers. Since customers cannot just "walk in" (as they would in a physical store), online businesses need to do a lot of advertising to get folks to visit their page—and since customers determine success or failure, advertising is critical to online businesses.

Lots of businesses made that their business model: instead of selling a product to customers, they would sell eyeballs to businesses that are trying to sell products. This leads to businesses giving away 'free' services to get eyeballs that they could sell to the highest bidder.

There is often a direct relationship between advertising and revenue: more advertising drives more revenue.

Many big online brands, including Google, Facebook, and to a large extent, Amazon, are in the business of funneling eyeballs to other products or services that generate revenue.

This trend also leads to analytics, and customer analysis. Big tech corps analyze their eyeballs in all conceivable ways to target them at specific customers.

4 Competition

The internet makes the world a smaller place. A customer can find a company anywhere int he world. Competition can find the cheapest place to operate, and compete on cost.

Lower cost to enter markets also means there are millions of online businesses selling anything imaginable. This often leads to fraud, causing customers to trust only a handful of online businesses—leading to a cycle where big companies get bigger, simply because everyone trusts them and does business with them.

5 Aggregators

Another kind of business is aggressors. These includes websites that compare prices across many different platforms—sending consumers to the cheapest for essentially the same service.

For example, if you need a flight from point A to point B, you can find the cheapest tickets via such an aggregator, irrelevant of what airlines may be publishing as a ticket price.

Often such aggregators have special deals that allow them to get better prices than going direct (you call up a hotel, get a quote of $200 per night, but book the same hotel through hotwire for $140).

6 Connectivity & Business Relationships

It is often beneficial for businesses to work together. For example, a retail store needs inventory to sell, and the supplier has a vested interest to keep those inventory numbers at proper levels. The retail store may let a supplier manage their inventory levels—automatically ordering, shipping and delivering more products whenever inventory levels fall below certain levels set by supplier.