Principles of Software Construction: Objects, Design, and Concurrency

Case Studies in Data Consistency and Google's PageRank

Charlie Garrod    Christian Kästner
Administrivia

• Homework 6, homework 6, homework 6...
  § Due Thursday, 11:59 p.m.
  § May turn in as late as Saturday, 11:59 p.m.

• Final exam review session
  § Saturday, May 10th, 6 – 8 p.m., PH 100

• Final exam
  § Monday, May 12th, 5:30 – 8:30 p.m., UC McConomy

• Faculty course evaluations
  § https://cmu.smartevals.com/

• TA feedback(?)
  § Email from Greg Kesden coming soon(?)
Last time...
Data consistency

• Suppose $D$ is the database for some application and $\varphi$ is a function from database states to \{true, false\}
  ▪ We call $\varphi$ an \textit{integrity constraint} for the application if $\varphi(D)$ is true if the state $D$ is "good"
  ▪ We say a database state $D$ is \textit{consistent} if $\varphi(D)$ is true for all integrity constraints $\varphi$
  ▪ We say $D$ is inconsistent if $\varphi(D)$ is false for any integrity constraint $\varphi$

• Transaction ACID properties:
  ▪ Atomicity: All or nothing
  ▪ Consistency: Application-dependent as before
  ▪ Isolation: Each transaction runs as if alone
  ▪ Durability: Database will not abort or undo work of a transaction after it confirms the commit
The CAP theorem for distributed systems

• For any distributed system you want...
  ▪ Consistency
  ▪ Availability
  ▪ tolerance of network Partitions

• ...but you can support at most two of the three
Today: Case study in consistency, and PageRank

- Google's PageRank algorithm
- Ruminations on data consistency

The PageRank Citation Ranking: Bringing Order to the Web

January 29, 1998

Abstract

The importance of a Web page is an inherently subjective matter, which depends on the reader's interests, knowledge, and attitude. There is no much that can be said objectively about the relative importance of Web pages. This paper describes PageRank, a method for using Web pages objectively and mechanically, effectively measuring the various interest and attention devoted to them.

We compare PageRank to an idealized random Web surfer. We show how to efficiently compute PageRank for large numbers of pages. And we show how to apply PageRank to search and to user navigation.

1 Introduction and Motivation

The World Wide Web creates many new challenges for information retrieval. It is very large and heterogeneous. Current estimates are that there are over 150 million web pages with a doubling life of less than one year. More importantly, the web pages are extremely diverse, ranging from "What is Joe having for lunch today?" to journals about information retrieval. In addition to these
A "university" search, circa 1997

From Page et al., "The PageRank Citation Ranking: Bringing Order to the Web"
Traditional information retrieval

- 1997’s http://www.net.cmu.edu:

  <TITLE>Carnegie Mellon University - Computing Services - Network Group</TITLE>


  <P>

  <H2>Departments</H2>

  <DL>
    <DD> <IMG SRC="http://www.net.cmu.edu/icons/greenball.gif"> <A HREF="http://www.net.cmu.edu/datacomm/home.html"> <B> Data Communications</B></A>
    ...
  </DL>
Improving IR with citation counts

• If a page is important, other pages link to it.

\[ r(v) = \sum_{(u,v) \in E} 1 \]
PageRank: weighted citations

- If a page is important, other important pages link to it.
PageRank: weighted citations

- If a page is important, other important pages link to it.

\[ r(v) = \sum_{(u,v) \in E} \frac{r(u)}{|\text{out-deg}(u)|} \]

- e.g.,

![Diagram of a network with nodes and edges labeled 1, 2, 3, 4, 6, v1, v2, v3, v4. Edges connect 1 to 2, 2 to 3, 3 to 4, 4 to 6, and 6 to v2.]
PageRank: weighted citations

• If a page is important, other important pages link to it

\[ r(v) = \sum_{(u,v) \in E} \frac{r(u)}{|\text{out-deg}(u)|} \]

- e.g.,
PageRank: weighted citations

- If a page is important, other important pages link to it

\[ r(v) = \sum_{(u,v) \in E} \frac{r(u)}{|\text{out-deg}(u)|} \]

- e.g.,
PageRank: weighted citations

- If a page is important, other important pages link to it.

\[ r(v) = \sum_{(u,v) \in E} \frac{r(u)}{|\text{out-deg}(u)|} \]

- Is this well-defined?
- How do we compute it?
- How do we compute it efficiently?
The WWW as a graph as a matrix

\[ W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} \]
The WWW as a graph as a matrix

- PageRanks $\mathbf{R} = [r_1, r_2, \ldots, r_n]$ solve the linear equation $\mathbf{R} = \mathbf{R} \cdot \mathbf{W}$
  - $\mathbf{R}$ is an eigenvector of the Web

$$\mathbf{W} = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 \\
0 & 1 & 0 & 0 \\
1/3 & 1/3 & 1/3 & 0 \\
\end{pmatrix}$$
The power method

• (under some conditions) To find an eigenvector $\mathbf{v}$ of a matrix $\mathbf{M}$
  - Start with some approximation of $\mathbf{v}$: $\mathbf{v}_0$
  - Compute repeatedly:

$$v_{i+1} = \frac{v_i \cdot M}{||v_i \cdot M||}$$
The power method for PageRank

- Assign some initial PageRank $R$
- While $R$ hasn't converged, compute “next” PageRanks from the previous PageRanks

PageRank(G,delta)

Initialize $R = $ something, $R' = 0$
while ($R - R' > delta$)

$R' = R$

$R = 0$

for each edge (u,v) in G

$R[v] += (R'[u] / \text{out-deg}(u))$
A PageRank example

\[
\begin{align*}
&v_1 \\
&v_2 \\
&v_3 \\
&v_4
\end{align*}
\]

\[
\begin{align*}
&1 \\
&1/2 \\
&1/3 \\
&1/3
\end{align*}
\]

\[
\begin{align*}
&1 \\
&1/2 \\
&1/3 \\
&1/3
\end{align*}
\]

\[
\begin{align*}
&v_1 & v_2 & v_3 & v_4 \\
R_0 & 0.25 & 0.25 & 0.25 & 0.25 \\
R_1 & 0.083333 & 0.583333 & 0.208333 & 0.125 \\
R_2 & 0.041666 & 0.333333 & 0.333333 & 0.291666 \\
R_3 & 0.097222 & 0.472222 & 0.263888 & 0.166666 \\
R_4 & 0.055555 & 0.416666 & 0.291666 & 0.236111 \\
R_5 & 0.078703 & 0.425925 & 0.287037 & 0.208333
\end{align*}
\]
Convergence of the power method

Theorem:

For any initial PageRanks summing to 1, the power method will converge to a well-defined, unique solution if the transition matrix $W$ is stochastic, aperiodic, and irreducible.
A stochastic transition matrix

- A transition matrix is *stochastic* if all rows sum to 1

\[
W = \begin{pmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 \\
0 & 1 & 0 & 0 \\
1/3 & 1/3 & 1/3 & 0
\end{pmatrix}
\]
A stochastic transition matrix

- A transition matrix is *stochastic* if all rows sum to 1
A stochastic transition matrix

- A transition matrix is *stochastic* if all rows sum to 1

\[
\begin{pmatrix}
0 & 1/3 & 1/3 & 1/3 \\
0 & 0 & 1/2 & 1/2 \\
0 & 1 & 0 & 0 \\
1/3 & 1/3 & 1/3 & 0 \\
\end{pmatrix}
\]
An aperiodic transition matrix

- A transition matrix is *periodic* if there is an integer $k > 1$ such that the interval between visits of two vertices is always a multiple of $k$.
An aperiodic transition matrix

- A transition matrix is *periodic* if there is an integer $k > 1$ such that the interval between visits of a vertex is always a multiple of $k$.
• The transition matrix is irreducible if it’s possible to (eventually) reach each state from any other state
An irreducible transition matrix

• The transition matrix is irreducible if it’s possible to (eventually) reach each state from any other state
Computing PageRank efficiently

- Can keep Web graph on disk
  - PageRanks in RAM
  - Do not store modifications that made $W$ stochastic, aperiodic, and irreducible
  - Use smart initial PageRanks

- Can partition Web graph between computers
Aside: Problems with PageRank

Biography of President George W. Bush
Biography of the president from the official White House web site.
www.whitehouse.gov/president/gwbio.html - 23k - Cached - Similar pages
Past Presidents - Kids Only - Current News - President
More results from www.whitehouse.gov »

Welcome to MichaelMoore.com!
Official site of the gadfly of corporations, creator of the film Roger and Me and the television show The Awful Truth. Includes mailing list, message board.
www.michaelmoore.com - 35k - Sep 1, 2005 - Cached - Similar pages

BBC NEWS | Americas | 'Miserable failure' links to Bush
Web users manipulate a popular search engine so an unflattering description linked to the president’s page.
news.bbc.co.uk/2/hi/americas/3298443.stm - 31k - Cached - Similar pages

Harmonic Cube Cubic Spirits Are Supreme Being

Welcome to the site where
RankMambo is Above Von Ahn

Luis von Ahn is Ignorance
You maybe academically retarded.

Academia Retards By Fact Google Has 1 Rank When Dead Still, And 4 Ranks Within 1 Google Crawling, losing 3 ranks retards humanity.

http://www.rankmambo.com/
Problem with PageRank computation…

- In spring 2000, Google's web-crawling system failed too frequently to update their web index
  - Their solution: Google File System and MapReduce
Problem with PageRank computation...

- In spring 2000, Google's web-crawling system failed too frequently to update their web index
  - Their solution: Google File System and MapReduce
- How bad is this web service outage?
  - ...in terms of data consistency
Data consistency at Facebook

• Replication for scalability:
  - Read-any, write-all
  - Palo Alto, CA is primary replica

  Aside: A 2010 conversation:
  Academic researcher: What would happen if X occurred?
  Facebook engineer: We don't know. X hasn't happened yet but it would be bad.
Data consistency at Amazon

- Strict data consistency increases real costs

Amazon engineer: "'Usually ships in 2-3 days'? What does that mean? Absolutely nothing."
A common reality: Relaxed data consistency

- Relaxed in time
  - E.g., Time-to-live in a data cache

- Relaxed in value
  - I.e., within some error bound from the correct value

- Other consistency guarantees
  - E.g., Causal consistency


Don't Settle for Eventual:
Scalable Causal Consistency for Wide-Area Storage with COPS

Wyatt Lloyd*, Michael J. Freedman*, Michael Kaminsky*, and David G. Andersen†
*Princeton University, †Intel Labs, ‡Carnegie Mellon University

ABSTRACT

Geo-replicated, distributed data stores that support complex online applications, such as social networks, must provide an "always-on" experience where operations always complete with low latency. Today’s systems often sacrifice strong consistency to achieve these goals, exposing inconsistencies to their clients and necessitating complex application logic. In this paper, we identify and define a consistency model—causal consistency with convergent conflict handling, or causal+—that is the strongest achieved under these constraints.

We present the design and implementation of COPS, a key-value store that delivers this consistency model across the wide-area. A key contribution of COPS is its scalability, which can enforce causal dependencies between keys stored across an entire cluster, rather than a single server like previous systems. The central approach in 1. INTRODUCTION

Distributed data stores are a fundamental building block of modern Internet services. Ideally, these data stores would be strongly consistent, always available for reads and writes, and able to continue operating during network partitions. The CAP Theorem, unfortunately, proves it impossible to create a system that achieves all three [13, 23]. Instead, modern web services have chosen overwhelmingly to embrace availability and partition tolerance at the cost of strong consistency [16, 20, 30]. This is perhaps not surprising, given that this choice also enables these systems to provide low latency for client operations and high scalability. Further, many of the earlier high-scale Internet services, typically focusing on web search, saw little reason for stronger consistency, although this position is changing with the rise of interactive services such as social networking applications [46]. We refer to systems with these four
Summary

• Google makes $billions by treating us all like random surfers
  ▪ PageRank as iterative, weighted citation rankings
  ▪ WWW graph modifications needed to compute PageRank

• Data consistency can be more than a boolean function
Thursday...

• Guest lecture by Claire Le Goues