Google’s PageRank

CS515 APPLICATION
DAVID F. GLEICH
A cartoon websearch primer

1. Crawl webpages
2. Analyze webpage text (information retrieval)
3. **Analyze webpage links**
4. Fit measures to human evaluations
5. Produce rankings
6. Continuously update
Gleich's syndrome

Gleich's syndrome or episodic angioedema with eosinophilia is a rare disease in which the body swells up episodically (angioedema), associated with raised antibodies of the IgM type and increased numbers of eosinophil granulocytes, a type of white blood cells, in the blood (eosinophilia). It was first described in 1984.[1]

Its cause is unknown, but it is unrelated to capillary leak syndrome (which may cause similar swelling episodes) and eosinophilia-myalgia syndrome (which features eosinophilia but alternative symptoms). Moreover, it is not a form of hypereosinophilic syndrome as there is no evidence that it leads to organ damage. Some studies have shown that edema attacks are associated with degranulation (release of enzymes and mediators from eosinophils), and others have demonstrated antibodies against endothelium (cells lining blood vessels) in the condition.[2]

Gleich syndrome has a good prognosis. Attack severity may improve with steroid treatment.[1][2]

Eosinophilia

Eosinophilia is the state of having a high concentration of eosinophils (eosinophil granulocytes) in the blood. The normal concentration is between 0 and 0.5 x 10^9 eosinophils per litre of blood. Eosinophilia can be reactive (in response to other stimuli such as allergy or infection) or non-reactive.

The release of interleukin 5 by T cells, mast cells and macrophages stimulates the production of eosinophils.

Causes

Hypereosinophilic syndrome

The hypereosinophilic syndrome (HS) is a disease characterized by a persistently elevated eosinophil count (≥ 1500 eosinophils/mm³) in the blood for at least six months without any recognizable cause, with involvement of either the heart, nervous system, or bone marrow.[1]

HS is a diagnosis of exclusion, after clonal eosinophilia (such as leukemia) and reactive eosinophilia (in response to infection, autoimmune disease, atopy, hypoadrenalism or cancer) have been ruled out.[2]

There are some associations with chronic...
The Model

1. follow edges uniformly with probability $\alpha$, and
2. randomly jump with probability $1 - \alpha$, we’ll assume everywhere is equally likely
The Model

1. follow edges uniformly with probability \( \frac{1}{6} \), and
2. randomly jump with probability \( \frac{1}{2} \), we’ll assume everywhere is equally likely

The places we find the surfer most often are important pages.

“jump” \( \rightarrow \) \( \mathbf{v} = \left[ \frac{1}{n} \ldots \frac{1}{n} \right]^T \)

\[
\begin{pmatrix}
\frac{1}{6} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
\frac{1}{6} & 0 & 0 & \frac{1}{3} & 0 & 0 \\
\frac{1}{6} & \frac{1}{2} & 0 & \frac{1}{3} & 0 & 0 \\
\frac{1}{6} & 0 & \frac{1}{2} & 0 & 0 & 0 \\
\frac{1}{6} & 0 & \frac{1}{2} & \frac{1}{3} & 0 & 1 \\
\frac{1}{6} & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}
\]

\( P \) is a Markov chain, and \( \mathbf{v}^T P = \mathbf{e}^T \)

\( P_{ij} \geq 0 \)

\( e^T P = e^T \)

\( \mathbf{e}^T \mathbf{v} = 1 \)
The CS515 definition of PageRank

A PageRank vector $\mathbf{x}$ is the solution of the linear system:

$$(I - \alpha \mathbf{P}) \mathbf{x} = (1 - \alpha) \mathbf{v}$$

where $\mathbf{P}$ is a column stochastic matrix, $0 \leq \alpha < 1$, and $\mathbf{v}$ is a probability vector.
Applies anywhere!

GeneRank ProteinRank FoodRank SportsRank HostRank TrustRank BadRank IsoRank SimRank ObjectRank ItemRank ArticleRank BookRank FutureRank TimedPageRank SocialPageRank DiffusionRank ImpressionRank TweetRank TwitterRank ReversePageRank PageTrust PopRank CiteRank FactRank InvestorRank ImageRank VisualRank QueryRank BookmarkRank StoryRank PerturbationRank ChemicalRank RoadRank PaperRank
Let’s show it always exists and how fast it converges.