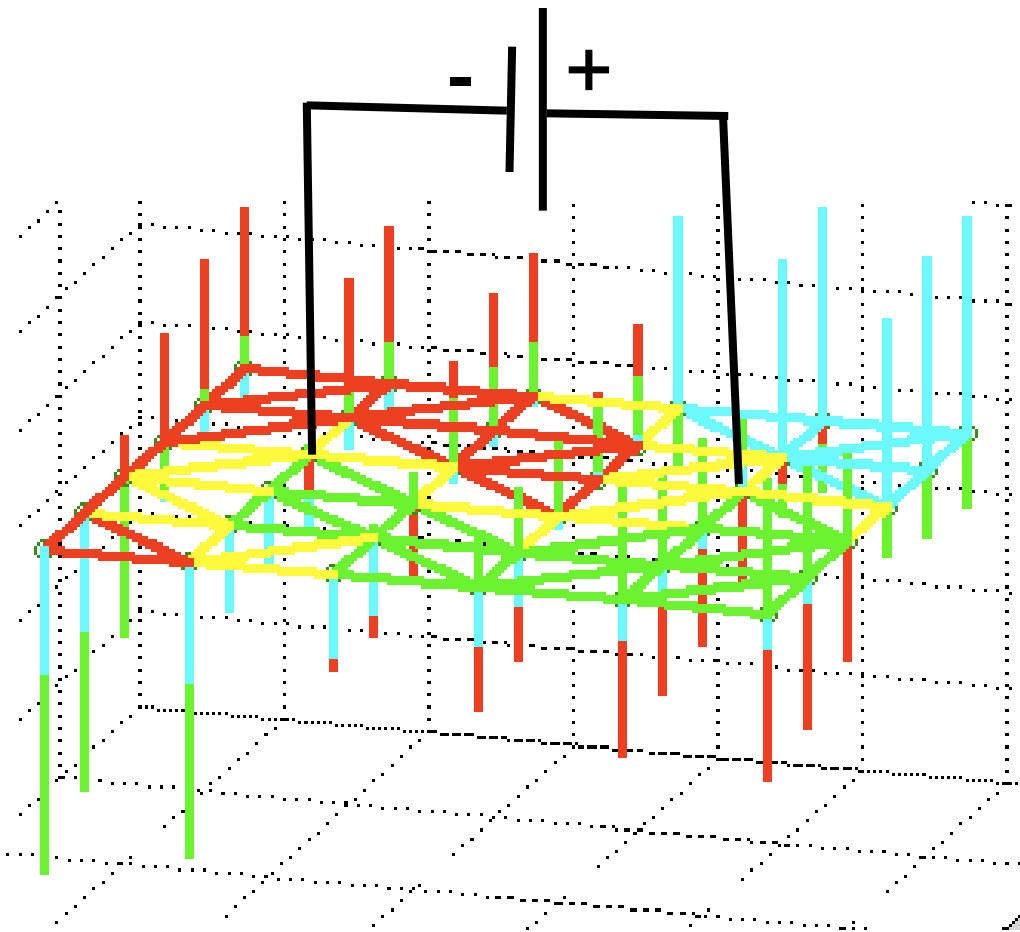


Unsupervised Learning in Complex Systems

Part C: Ether Dipsomania in Complex Systems*

CSSS 2008



Gregory Leibon
Memento Security & Dartmouth College

* or...Why one should NEVER analyze a complex system with out a pickle and 9 volt battery.

What we do...

- Introduce the Green's function on a Network.
- Produce the Green's embedding of the states where the Euclidean distances between the images are square root the commute distances in the network.
- Examine a applications to the Wiki Network an the Netflix's Problem.



Prove the mighty Commute Theorem!

MATHEMATICAL PAPERS

OF THE LATE

GEORGE GREEN,

FELLOW OF GONVILLE AND CAIUS COLLEGE, CAMBRIDGE.

AN ESSAY

ON THE APPLICATION OF MATHEMATICAL ANALYSIS

TO THE THEORIES

OF ELECTRICITY AND MAGNETISM.*

Circa 1828

A little deeper...

Green's Function

$$G(x, y)$$

Potential at x for point source at y .

$$\Delta G(x, y) = \delta y$$

(1.) THE function which represents the sum of all the electric particles acting on a given point divided by their respective distances from this point, has the property of giving, in a very simple form, the forces by which it is solicited, arising from the whole electrified mass.—We shall, in what follows, endeavour to discover some relations between this function, and the density of the electricity in the mass or masses producing it, and apply the relations thus obtained, to the theory of electricity.

Firstly, let us consider a body of any form whatever, through which the electricity is distributed according to any given law, and fixed there, and let x', y', z' , be the rectangular co-ordinates of a particle of this body, ρ' the density of the electricity in this particle, so that $dx'dy'dz'$ being the volume of the particle, $\rho'dx'dy'dz'$ shall be the quantity of electricity it contains: moreover, let r' be the distance between this particle and a point p exterior to the body, and V represent the sum of all the particles of electricity divided by their respective distances from this point, whose co-ordinates are supposed to be x, y, z , then shall we have

$$r' = \sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2},$$

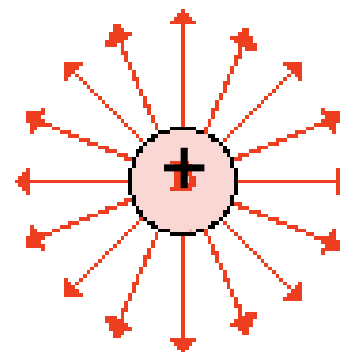
and

$$V = \int \frac{\rho' dx' dy' dz'}{r'};$$

the integral comprehending every particle in the electrified mass under consideration.

Example (R^3):

$$G(x, y) = \frac{1}{4\pi |x - y|}$$



....with second
source at
infinity

LAPLACE has shown, in his *Méc. Céleste*, that the function V has the property of satisfying the equation

$$0 = \frac{d^2 V}{dx^2} + \frac{d^2 V}{dy^2} + \frac{d^2 V}{dz^2},$$

and as this equation will be incessantly recurring in what follows, we shall write it in the abridged form $0 = \delta V$; the symbol δ being used in no other sense throughout the whole of this Essay.

In order to prove that $0 = \delta V$, we have only to remark, that by differentiation we immediately obtain $0 = \delta \frac{1}{r}$, and consequently each element of V substituted for V in the above equation satisfies it; hence the whole integral (being considered as the sum of all these elements) will also satisfy it. This reasoning ceases to hold good when the point p is within the body, for then, the coefficients of some of the elements which enter into V becoming infinite, it does not therefore necessarily follow that V satisfies the equation

$$0 = \delta V,$$

although each of its elements, considered separately, may do so.

...

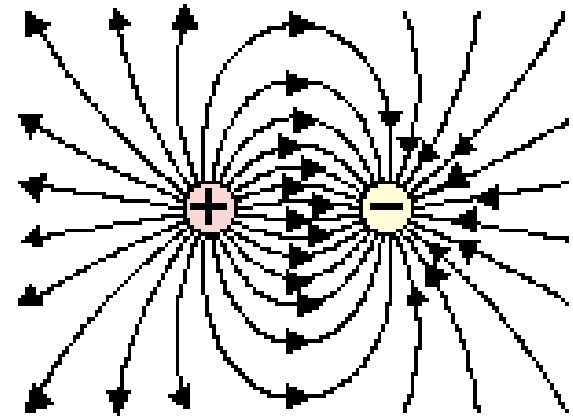
Hence, throughout the interior of the mass

$$0 = \delta V + 4\pi\rho;$$

of which, the equation $0 = \delta V$ for any point exterior to the body is a particular case, seeing that, here $\rho = 0$.

$$G(x, y_1, y_2)$$

$$G(x, y_1) - G(x, y_2)$$



$$\langle \Delta f, c \rangle = \langle f, \Delta c \rangle = 0$$

Searching for f such that

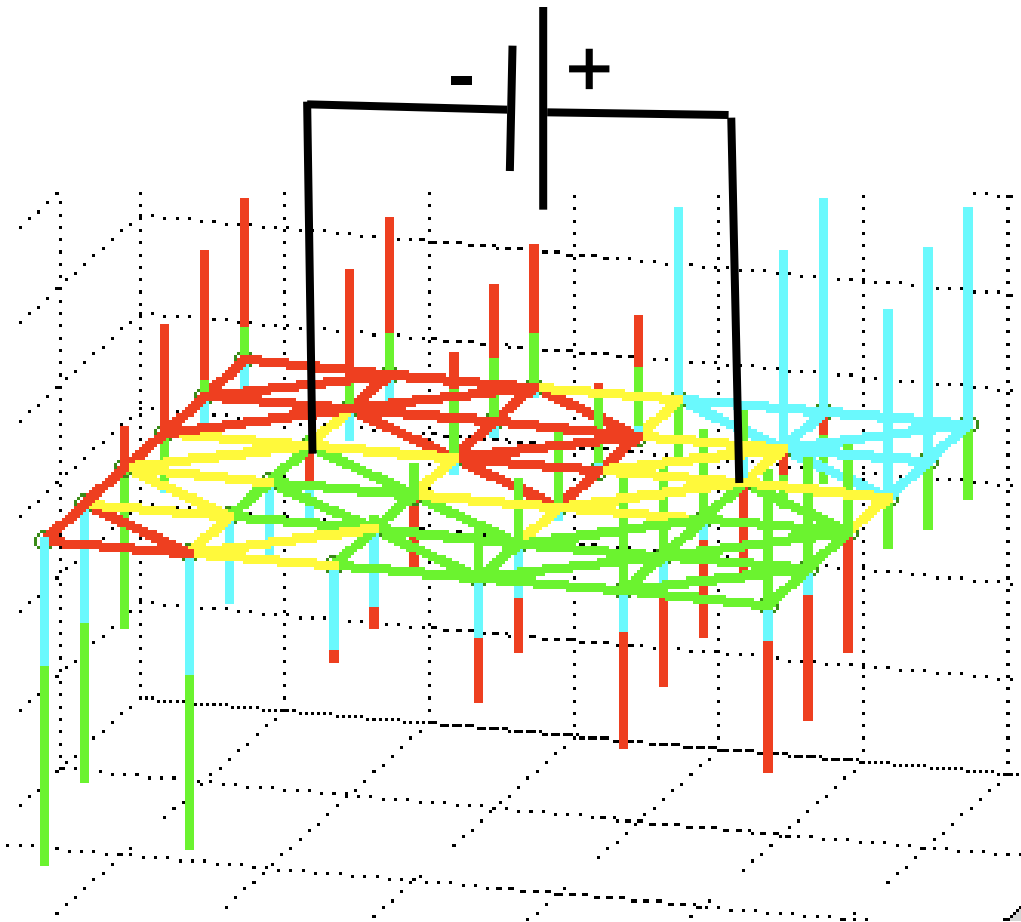
$$\Delta f = \frac{e_j}{\mu_j} - \frac{e_i}{\mu_i} = \mu^{\cdot -1}(e_j - e_i)$$

where

$$e_j(k) = \begin{cases} 1 & k = j \\ 0 & \text{otherwise} \end{cases}$$

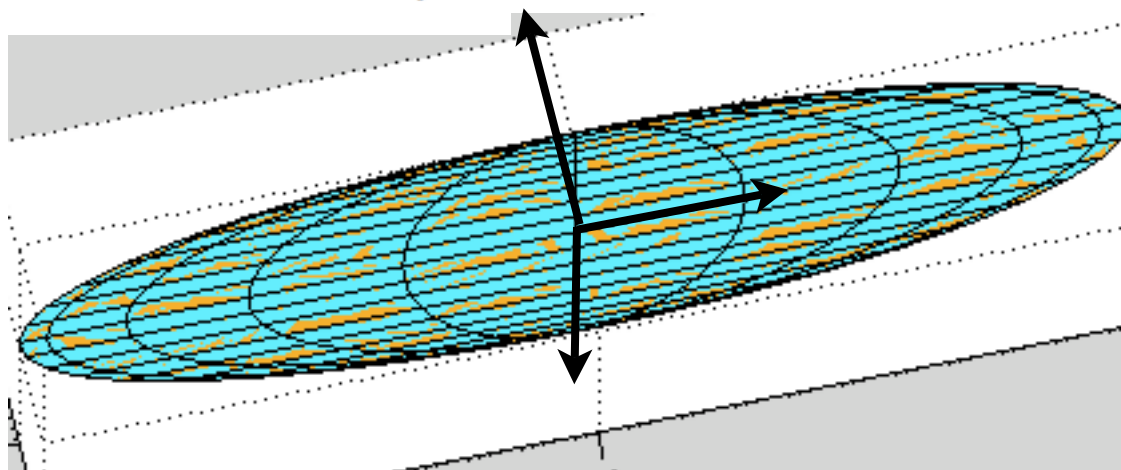
or

$$f = \Delta^{-1} \left(\mu^{\cdot -1}(e_j - e_i) \right)$$



$$f = \Delta^{-1} (\mu^{-1}(e_j - e_i))$$

Moore-Penrose
Pseudo Inverse



$$A = ODO^{tr}$$

$$f(A) = Of(D)O^{tr}$$

$$x^{-p} = \begin{cases} \bar{x}^p, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Spectral
Calculus

$$A = O^{tr} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} O$$

$$f(A) = O^{tr} \begin{bmatrix} f(0) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{bmatrix} O$$

Embedding Theorem $\left\| \Delta^{-1} \left(\mu^{-1} (e_j - e_i) \right) \right\|_{Dir} = \| G(e_j - e_i) \|_{Euclid}$

$$G = \mu^{\frac{1}{2}} \Delta^{-\frac{1}{2}} \mu^{-1}$$

So G embeds the states in \mathbb{R}^N such that the Euclidean distance squared is the Green's Distance.

Proof: $\langle \Delta^{-1} \mu^{-1} g, \Delta^{-1} \mu^{-1} h \rangle_{Dir} = \langle \Delta \Delta^{-1} \mu^{-1} g, \Delta^{-1} \mu^{-1} h \rangle$ definition

$$= \langle \Delta^{\frac{1}{2}} \Delta^{-1} \mu^{-1} g, \Delta^{\frac{1}{2}} \Delta^{-1} \mu^{-1} h \rangle$$

$$= \langle \Delta^{-\frac{1}{2}} \mu^{-1} g, \Delta^{-\frac{1}{2}} \mu^{-1} h \rangle$$

$$= (\Delta^{-\frac{1}{2}} \mu^{-1} g)^{tr} \mu^{\frac{1}{2}} \Delta^{-\frac{1}{2}} \mu^{-1} h$$

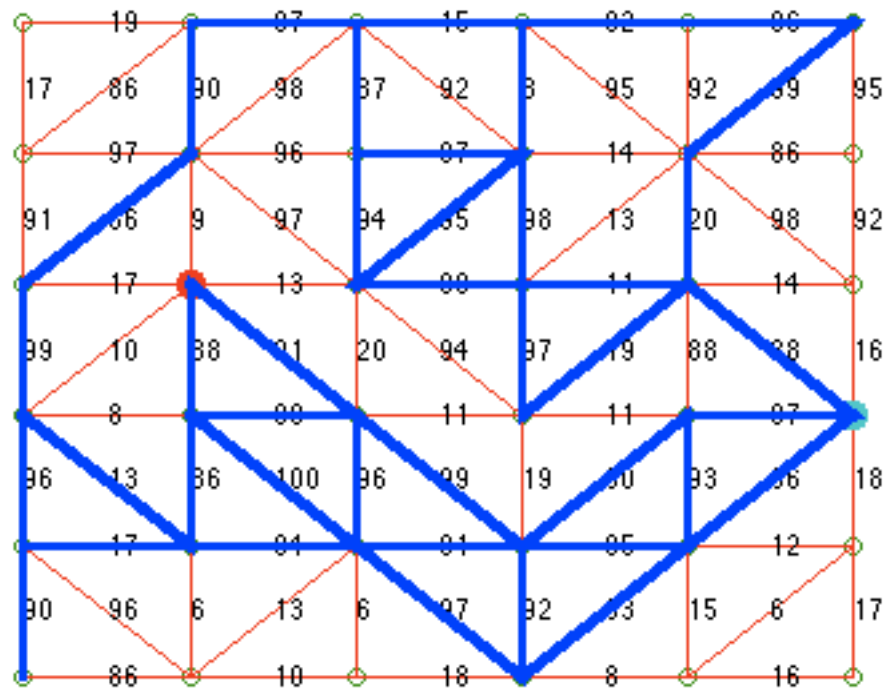
$$= (\mu^{\frac{1}{2}} \Delta^{-\frac{1}{2}} \mu^{-1} g)^{tr} \mu^{\frac{1}{2}} \Delta^{-\frac{1}{2}} \mu^{-1} h$$

$$= \langle Gg, Gh \rangle_{Euclid}$$

self adjoint
spectral calculus
definition
re-organize
definition

Definition: The Dirichlet Inner Product: $\langle g, h \rangle_{Dir} = \langle \Delta g, h \rangle$

The Commute Time from i to j (denoted C_{ij}) is the expected number up step needed to get from i to j and back to i .



116

```
[states Time]=Commute(15,25,1000);
```

Commute Theorem:

$$\left\| \Delta^{-1} \left(\mu^{-1} (e_j - e_i) \right) \right\|_{Dir}^2 = C_{ij}$$

Gauss Bonnet Theorem, Selberg trace Formula, Riemann Mapping Theorem, Commute Theorem, Classification of Surfaces,

P.G. Doyle and J. L. Snell. Random Walks and Electric Networks. The Mathematical Association of America, 1984.

A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari. The electrical resistance of a graph captures its commute and cover times. *Symposium on Theory of Computing*, pages 574–586, Seattle, WA, May 1989. <http://citeseer.ist.psu.edu/chandra89electrical.html> [More](#)

Klein, D. and Randic, M. (1993). Resistance distance. *Journal of Mathematical Chemistry*, 12, 81-95.

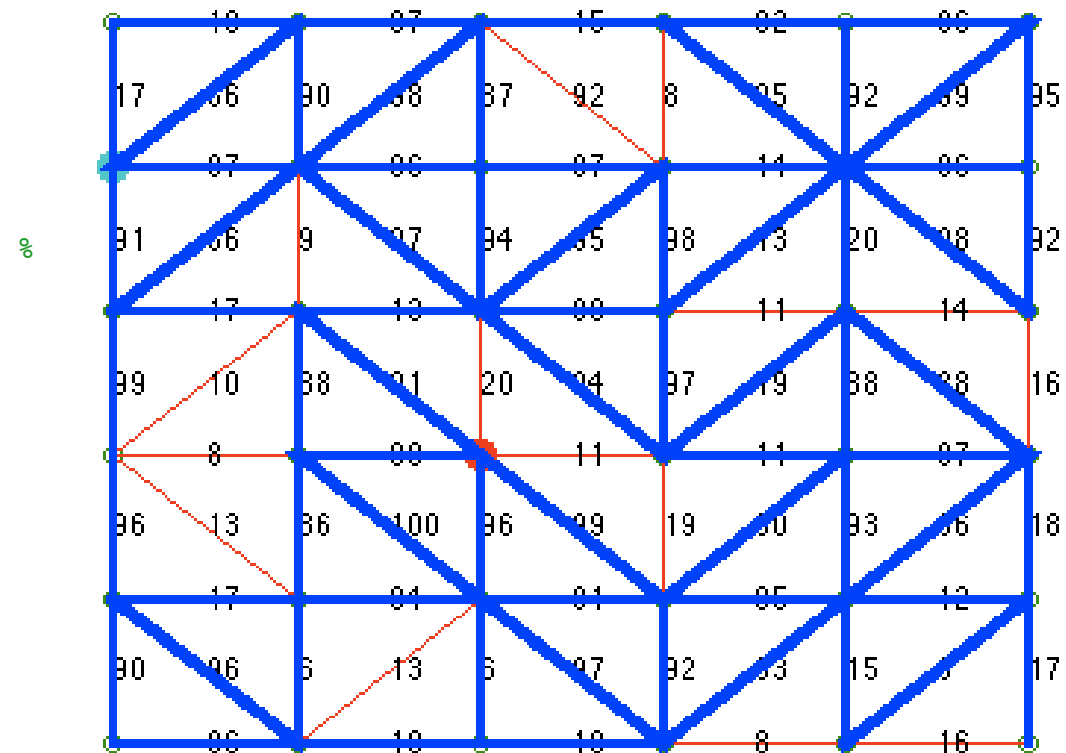
Francois Fouss, Alain Pirotte, Jean-Michel Renders, Marco Saerens: Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation. *IEEE Trans. Knowl. Data Eng.* 19(3): 355-369 (2007)

In MatLab

ECom(15,25)

ans =

140.8849



Steps383 Returns17

Commute(15,25,1000);

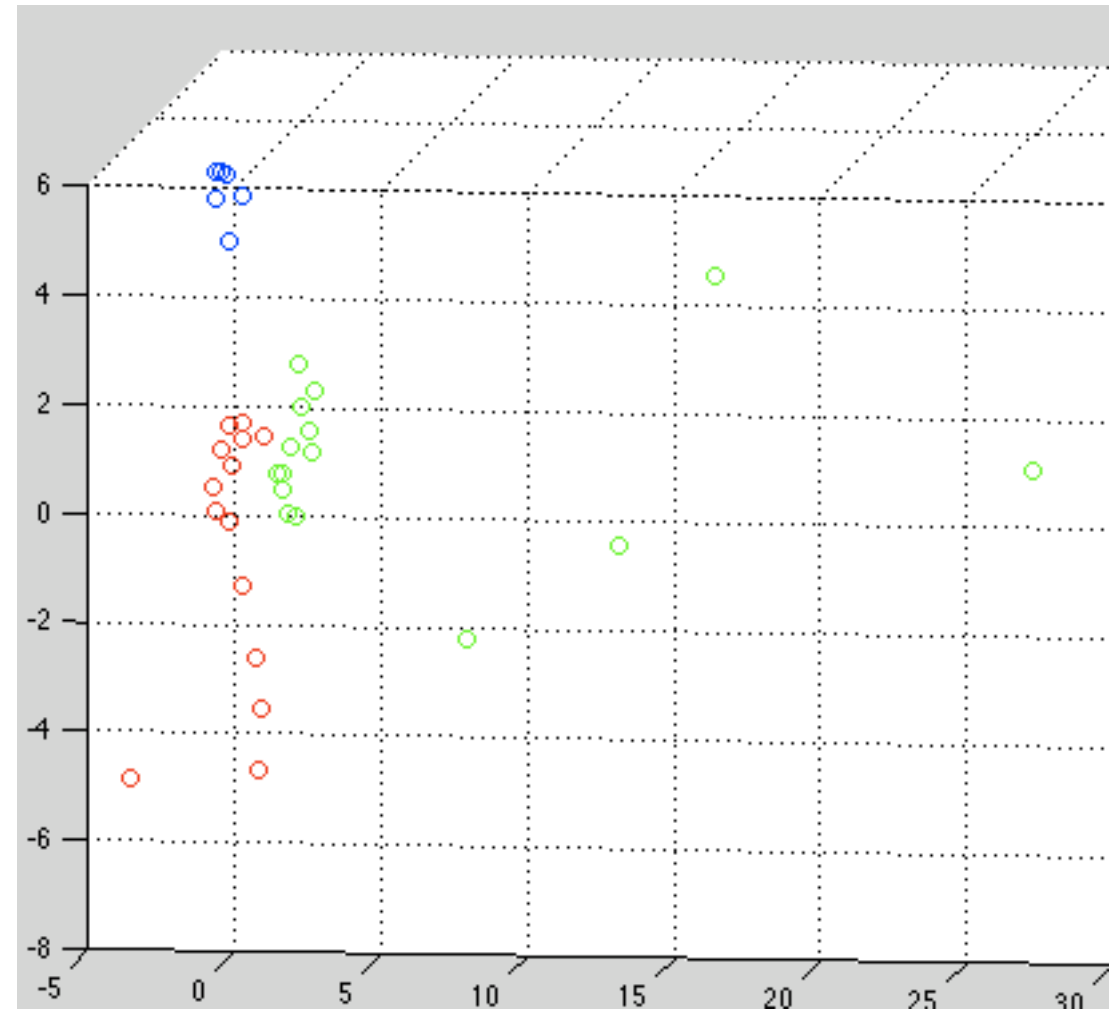
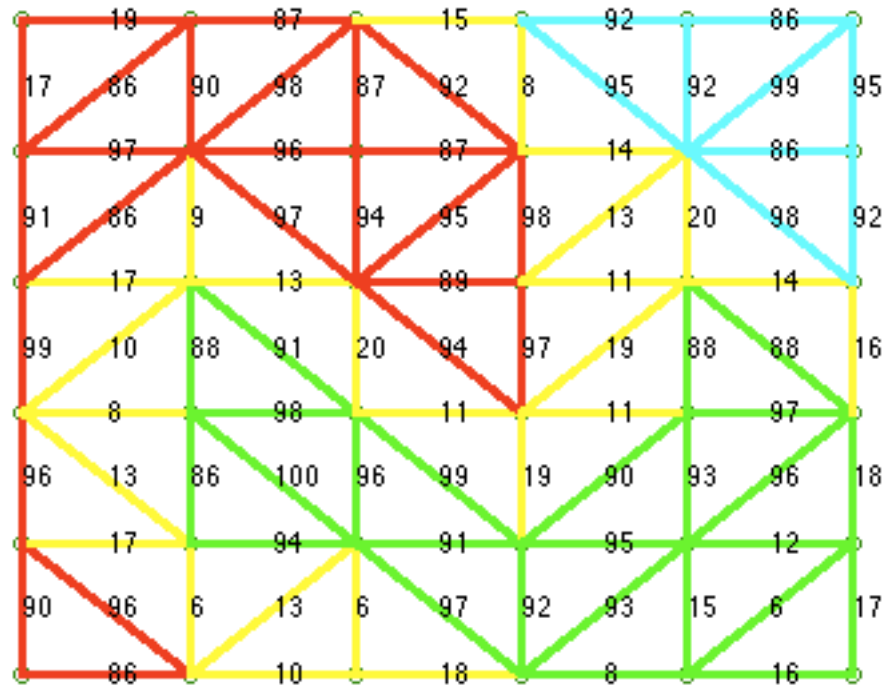
```
function Cij=ECom(i,j)
```

```
% This function computes the expected time  
% to commute from i to j and back to i  
% Example: i=15; j=25; ECom(i,j)
```

```
% We precomputed the transtion matrix T  
load T.mat; load mu.mat;  
% Compute Laplacian  
Lap=diag(ones(length(T),1)) - T;  
% Compute Laplacian's Psudo Inverse  
LapInv=pinv(Lap);  
% Construct the Charge Vector  
eij=zeros(size(mu));  
eij(i)=1;  
eij(j)=-1;  
charge=(1./mu).*eij;  
% Solve Laplacian f= Charge  
f=(LapInv*(charge)');  
% Compute the needed Dirichlet norm  
Cij=f'*diag(mu)*Lap*f;
```

```
end
```

MDS the Green's embedding to 3 dimensions



The Netflix Challenge

NETFLIX

Netflix Prize

Home Rules Leaderboard Register Update Submit Download

NETFLIX

Browse Recommendations Friends Queue Buy DVDs

Home Genres New Releases Previews Netflix Top 100 Crit

Movies For You

Randy, the following movies were chosen based on your interest in:
[Bowling for Columbine](#)
[Carnivale: Season 1](#)
[Fahrenheit 9/11](#)

You really liked it...
Now own it for just \$5.99
Shop as low as \$5.99

Welcome!

The Netflix Prize seeks to substantially improve the accuracy of predictions about how much someone is going to love a movie based on their movie preferences. Improve it enough and you win one (or more) Prizes. Winning the Netflix Prize improves our ability to connect people to the movies they love.

Read the [Rules](#) to see what is required to win the Prizes. If you are interested in joining the quest, you should [register a team](#).

You should also read the [frequently-asked questions](#) about the Prize. And check out how various teams are doing on the [Leaderboard](#).

Good luck and thanks for helping!

Guides:
Member Favorites
Easter Eggs
By Decade
By Studio
Movies You've Seen

NETFLIX

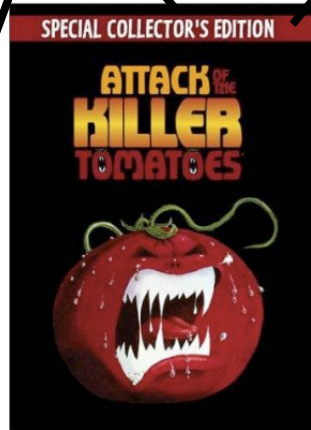
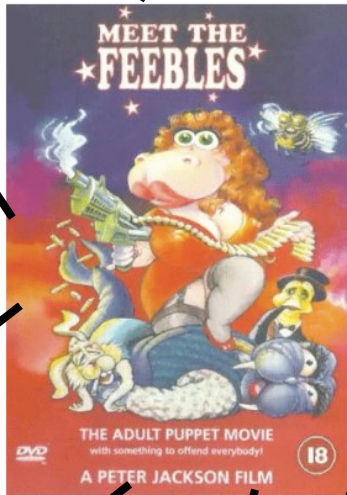
Netflix Prize

Home Rules Leaderboard Register Update Submit Download

Leaderboard

Display top 40 leaders.

Rank	Team Name	Best Score	% Improvement	Last Submitted
—	No Grand Prize candidates yet	—	—	—
Grand Prize - RMSE <= 0.8563				
—	No Progress Prize candidates yet	—	—	—
Progress Prize - RMSE <= 0.8625				
1	BellKor	0.8643	9.15	2008-05-30 13:00
2	BigChaos	0.8672	8.85	2008-06-07 17:00
3	When Gravity and Dinosaurs Unite	0.8675	8.82	2008-05-09 13:00
4	Gravity	0.8687	8.69	2008-06-06 20:00
5	Just a guy in a garage	0.8705	8.50	2008-06-06 14:00
6	acmehill	0.8709	8.46	2008-06-06 08:00
Progress Prize 2007 - RMSE = 0.8712 - Winning Team: KorBell				
7	KorBell	0.8712	8.43	2007-10-01 23:00
8	basho	0.8714	8.41	2008-05-21 22:00
9	PragmaticTheory	0.8721	8.34	2008-06-04 01:00
10	Dan Tillberg	0.8723	8.31	2008-03-28 20:00
11	Ces	0.8745	8.08	2008-04-24 03:00
12	Reel Ingenuity	0.8747	8.06	2008-04-02 21:00
13	Dinosaur Planet	0.8753	8.00	2007-10-04 04:00
14	OperaSolutions	0.8763	7.89	2008-06-06 16:00
15	pengpeng	0.8765	7.87	2008-06-06 17:00
16	JDennisSu	0.8768	7.84	2008-06-06 19:00
17	Efratko	0.8771	7.81	2008-05-29 16:00
18	Emily's Horse	0.8773	7.79	2008-05-31 16:00
19	Craig Carmichael	0.8777	7.75	2008-06-01 18:00
20	Three Blind Mice	0.8778	7.74	2008-02-16 20:00
21	JustAMan	0.8778	7.74	2008-05-26 16:00



Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation

François Fouss, Alain Pirotte, *Member, IEEE*, Jean-Michel Renders, and Marco Saerens, *Member, IEEE*

Abstract—This work presents a new perspective on characterizing the similarity between elements of a database or, more generally, nodes of a weighted and undirected graph. It is based on a Markov-chain model of random walk through the database. More precisely, we compute quantities (the **average commute time**, the **pseudoinverse of the Laplacian matrix** of the graph, etc.) that provide similarities between any pair of nodes, having the nice property of increasing when the number of paths connecting those elements increases and when the “length” of paths decreases. It turns out that the square root of the average commute time is a Euclidean distance and that the pseudoinverse of the Laplacian matrix is a kernel matrix (its elements are inner products closely related to commute times). A principal component analysis (PCA) of the graph is introduced for computing the subspace projection of the node vectors in a manner that preserves as much variance as possible in terms of the Euclidean commute-time distance. This graph PCA provides a nice interpretation to the “**Fiedler vector**,” widely used for graph partitioning. The model is evaluated on a collaborative-recommendation task where suggestions are made about which movies people should watch based upon what they watched in the past. Experimental results on the MovieLens database show that the Laplacian-based similarities perform well in comparison with other methods. The model, which nicely fits into the so-called “statistical relational learning” framework, could also be used to compute document or word similarities, and, more generally, it could be applied to machine-learning and pattern-recognition tasks involving a relational database.

6 EXPERIMENTS

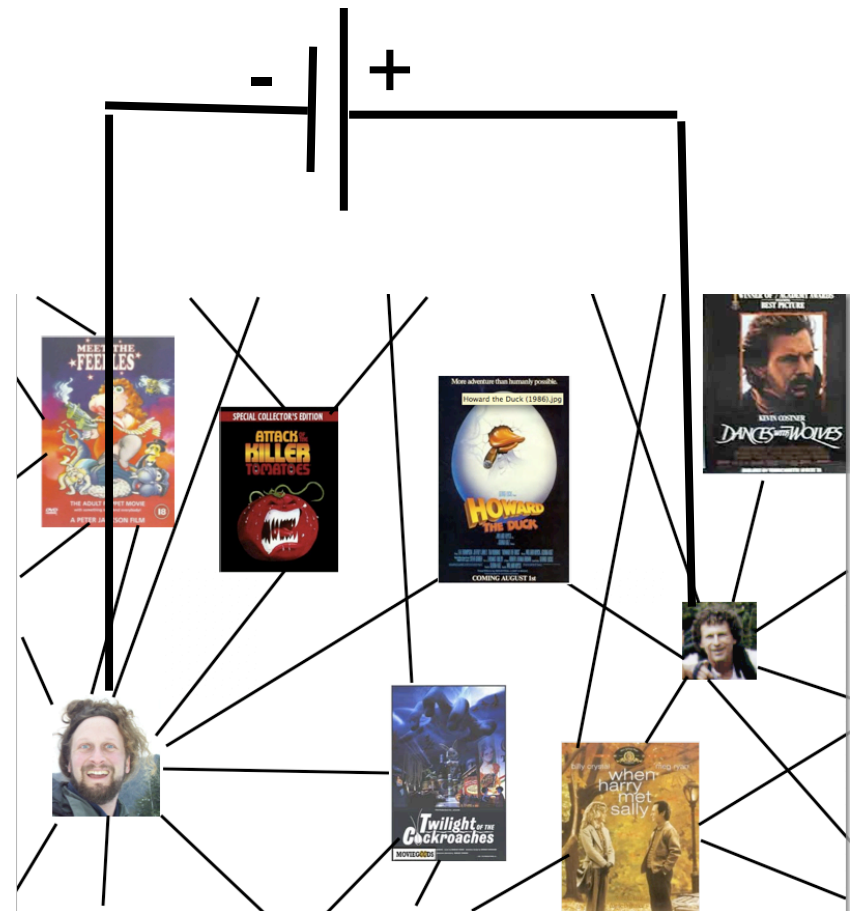
6.1 Experimental Methodology

Remember that each element of the people and the movie sets corresponds to a node of the graph. Each node of the people set is connected by a link to each movie watched by the corresponding person. Notice that, in this special case, the graph is bipartite. The results shown here do not take into account the numerical value of the ratings provided by the persons but only the fact that a person has or has not watched a movie (i.e., entries in the person-movie matrix are 0s and 1s). Moreover, our experiments do not take the movie_category set into account so that comparisons between the various scoring algorithms remain fair. Indeed, three standard scoring algorithms (i.e., maximum frequency, cosine, and nearest-neighbor algorithms) cannot naturally use the movie_category set to rank the movies.

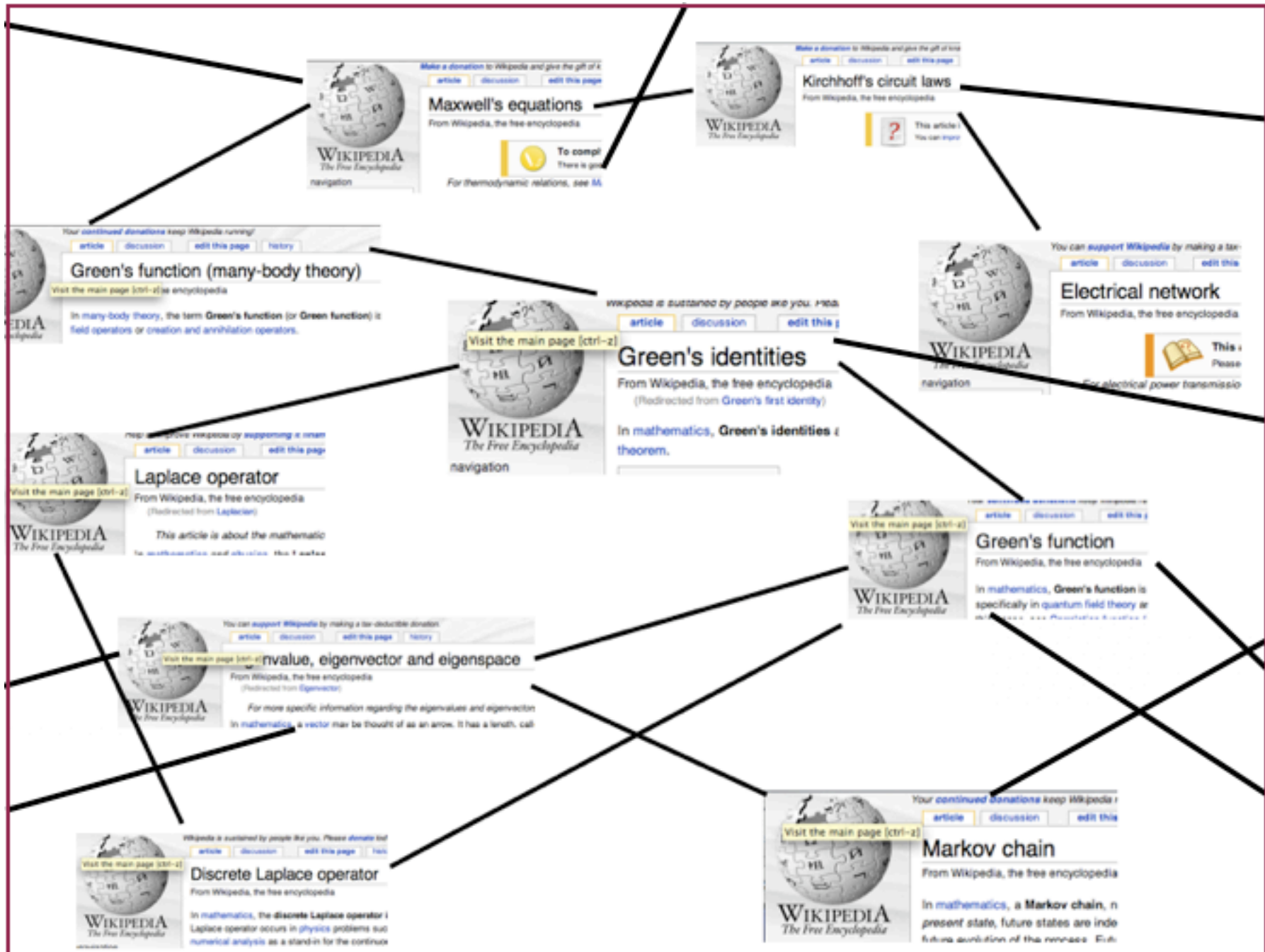
6.1.1 Data Set

Our experiments were performed on a real movie database from the Web-based recommender system MovieLens (<http://www.movielens.umn.edu>). We used a sample of their database as suggested in [59]: Enough people (i.e., 943 people) were randomly selected to obtain 100,000 ratings (considering only persons that had rated 20 or more movies on a total of 1,682 movies).

Great Project:
Do this for the
Netflix data!



A Wiki Adventure!



Abstract

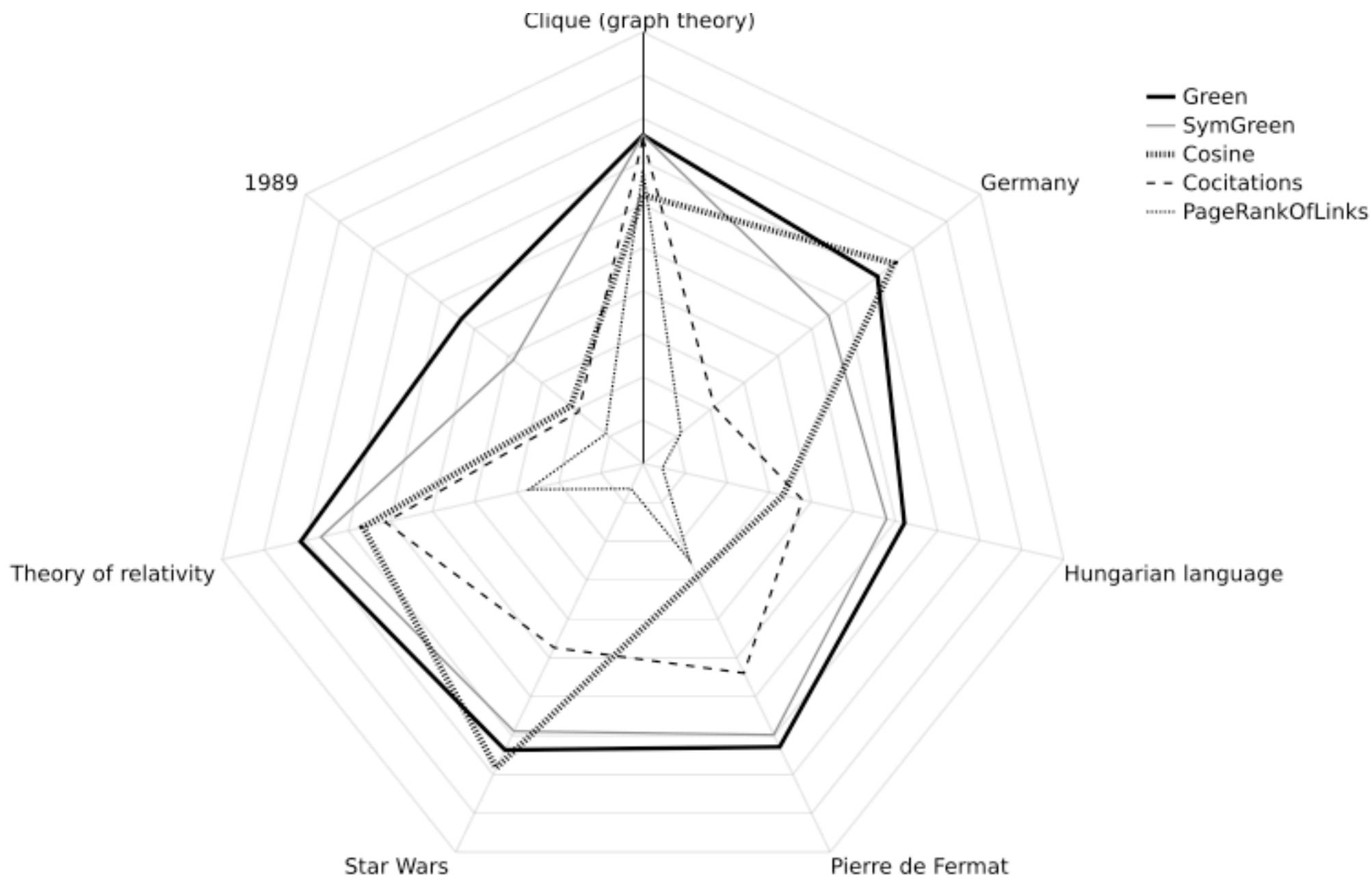
We introduce a new method for finding nodes semantically related to a given node in a hyperlinked graph: the Green method, based on a classical Markov chain tool. It is generic, adjustment-free and easy to implement. We test it in the case of the hyperlink structure of the English version of Wikipedia, the on-line encyclopedia. We present an extensive comparative study of the performance of our method versus several other classical methods in the case of Wikipedia. The Green method is found to have both the best average results and the best robustness.

Evaluation Methodology. We carried out a blind evaluation of the methods on 7 different articles, chosen for their diversity: (i) *Clique (graph theory)*: a very short, technical article. (ii) *Germany*: a very large article. (iii) *Hungarian language*: a medium-sized, quite technical article. (iv) *Pierre de Fermat*: a short biographical article. (v) *Star Wars*: a large article, with an important number of links. (vi) *Theory of relativity*: a short introductory article pointing to more specialized articles. (vii) *1989*: a very large article, containing all the important events of year 1989. It was unreasonable to expect our testers to evaluate more articles. In order to avoid any bias, we did not run the methods on these 7 articles before the evaluation procedure was launched.

People were asked to assign a mark between 0 and 10 (10 being the best) to the list of the first 20 results returned by

Table 1: Output of **GREEN** on the articles used for evaluation.

<i>Clique (graph theory)</i>	<i>Germany</i>	<i>Hungarian language</i>	<i>Pierre de Fermat</i>	<i>Star Wars</i>	<i>Theory of relativity</i>	<i>1989</i>
1. Clique (graph theory) 2. Graph (mathematics) 3. Graph theory 4. Category:Graph theory 5. NP-complete 6. Complement graph 7. Clique problem 8. Complete graph 9. Independent set 10. Maximum common subgraph isomorphism problem 11. Planar graph 12. Glossary of graph theory 13. Mathematics 14. Connectivity (graph theory) 15. Computer science 16. David S. Johnson 17. Independent set problem 18. Computational complexity theory 19. Set 20. Michael Garey	1. Germany 2. Berlin 3. German language 4. Christian Democratic Union (Germany) 5. Austria 6. Hamburg 7. German reunification 8. Social Democratic Party of Germany 9. German Empire 10. German Democratic Republic 11. Bavaria 12. Stuttgart 13. States of Germany 14. Munich 15. European Union 16. National Socialist German Workers Party 17. World War II 18. Jean Edward Smith 19. Soviet Union 20. Rhine	1. Hungarian language 2. Slovakia 3. Romania 4. Slovenia 5. Hungarian alphabet 6. Hungary 7. Croatia 8. Category:Hungarian language 9. Turkic languages 10. Finno-Ugric languages 11. Austria 12. Serbia 13. Uralic languages 14. Ukraine 15. Hungarian grammar (verbs) 16. German language 17. Hungarian grammar 18. Khanty language 19. Hungarian phonology 20. Finnish language	1. Pierre de Fermat 2. Toulouse 3. Fermat's Last Theorem 4. Diophantine equation 5. Fermat's little theorem 6. Fermat number 7. Grandes écoles 8. Blaise Pascal 9. France 10. Pseudoprime 11. Lagrange's four-square theorem 12. Number theory 13. Fermat polygonal number theorem 14. Holographic will 15. Diophantus 16. Euler's theorem 17. Pell's equation 18. Fermat's theorem on sums of two squares 19. Fermat's spiral 20. Fermat's factorization method	1. Star Wars 2. Dates in Star Wars 3. Palpatine 4. Jedi 5. Expanded Universe (Star Wars) 6. Star Wars Episode I: The Phantom Menace 7. Star Wars Episode IV: A New Hope 8. Obi-Wan Kenobi 9. Star Wars Episode III: Revenge of the Sith 10. Coruscant 11. Anakin Skywalker 12. Lando Calrissian 13. Luke Skywalker 14. Star Wars: Clone Wars 15. List of Star Wars books 16. George Lucas 17. Star Wars Episode II: Attack of the Clones 18. Splinter of the Mind's Eye 19. List of Star Wars comic books 20. The Force (Star Wars)	1. Theory of relativity 2. Special relativity 3. General relativity 4. Spacetime 5. Lorentz covariance 6. Albert Einstein 7. Principle of relativity 8. Electromagnetism 9. Lorentz transformation 10. Inertial frame of reference 11. Speed of light 12. Galilean transformation 13. Local symmetry 14. Category:Relativity 15. Galilean invariance 16. Gravitation 17. Global symmetry 18. Tensor 19. Maxwell's equations 20. Introduction to general relativity	1. 1989 2. Cold War 3. 1912 4. Tiananmen Square protests of 1989 5. Soviet Union 6. German Democratic Republic 7. George H. W. Bush 8. 1903 9. Communism 10. 1908 11. 1929 12. Ruhollah Khomeini 13. March 1 14. Czechoslovakia 15. June 4 16. The Satanic Verses (novel) 17. 1902 18. November 7 19. October 9 20. March 14
Mark: 7.6/10	Mark: 7.0/10	Mark: 6.2/10	Mark: 7.3/10	Mark: 7.4/10	Mark: 8.1/10	Mark: 5.4/10



Should I prove the Commute Theorem?!

Commute Theorem:

$$\left\| \Delta^{-1} \left(\mu^{-1} (e_j - e_i) \right) \right\|_{Dir}^2 = C_{ij}$$

Commutate Theorem:

$$\left\| \Delta^{-1} \left(\mu^{-1}(e_j - e_i) \right) \right\|_{Dir}^2 = C_{ij}$$

Proof: Let p_e be the probability of reaching j before i when starting at i .

Proposition 1 (Renewal Theory)

$$C_{ij} = \frac{1}{\mu_i p_e}$$

Proposition 2 (Maximum Principle, Potential Theory)

$$\left\| \Delta^{-1} \left(\mu^{-1}(e_j - e_i) \right) \right\|_{Dir}^2 = \frac{1}{\mu_i p_e}$$

Proposition I

$$C_{ij} = \frac{1}{\mu_i p_e}$$

$R = \#$ Steps taken going from i to j

$N = \#$ Returns to i on commute

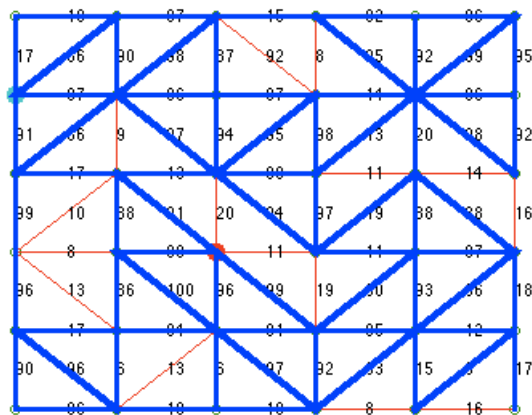
Second Fundamental

Mystery
of Probability
Theory

(see the appendix)

$$E(C_{ij}) = E\left(\sum_{k=1}^N R_k\right)$$

$$= E(N)E(R)$$

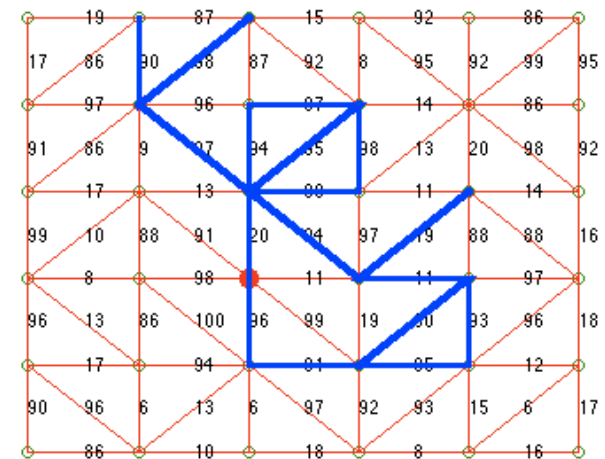


Steps 383 Returns 17

`Commute(15,25,1000);`

Very well
known!

$$E(R) = \frac{1}{\mu_i}$$



`Loop(15,1000);`

Proposition 2 $\left\| \Delta^{-1} \left(\mu^{-1} (e_j - e_i) \right) \right\|_{Dir}^2 = \frac{1}{\mu_i p_e}$

$$\Delta g(k) = 0 \quad k \neq i, j \quad \text{Dirichlet Problem}$$

$$g(i) = 0$$

$$g(j) = 1$$

Unique solution up to constant
from Maximum Principle



$g(k)$ equals the probability starting at k of reaching j before i .



$\Delta g(i)$ is minus the probability of escaping from i to j starting at i

$$\Delta g(j) = \frac{\mu_i p_e}{\mu_j}$$

constant function
in kernel

$$f = \frac{g}{u_i p_e}$$

where

$$f = \Delta^{-1} \left(\mu^{-1}(e_j - e_i) \right)$$

$$\|g\|_{Dir}^2 = \langle \Delta g, g \rangle$$

$$= \mu_j \frac{\mu_i p_e}{\mu_j} = \mu_i p_e$$

$$\left\| \Delta^{-1} \left(\mu^{-1}(e_j - e_i) \right) \right\|_{Dir}^2 = \frac{1}{(\mu_i p_e)^2} \|g\|_{Dir}^2 = \frac{1}{\mu_i p_e}$$

Poyla's Theorem

On the 2-d Euclidean lattice a drunk Lord will always find his way home. In three dimensions he may not be so lucky!

“I have not had a moment's peace or happiness in respect to electromagnetic theory since November 28, 1846. All this time I have been liable to fits of ether dipsomania, kept away at intervals only by rigorous abstention from thought on the subject.”

~Lord Kelvin~ (to FitzGerald 1896)

* **Dipsomania** is a term USUALLY related to an uncontrollable craving for alcohol.... the obsession is so compulsive that the dipsomaniac will ingest whatever intoxicating liquid is at hand, whether it is fit for consumption or not. Dipsomania differs from [alcoholism](#) in that it is an uncontrollable periodic lust for alcohol, with, in the interim, no desire for alcoholic beverages.

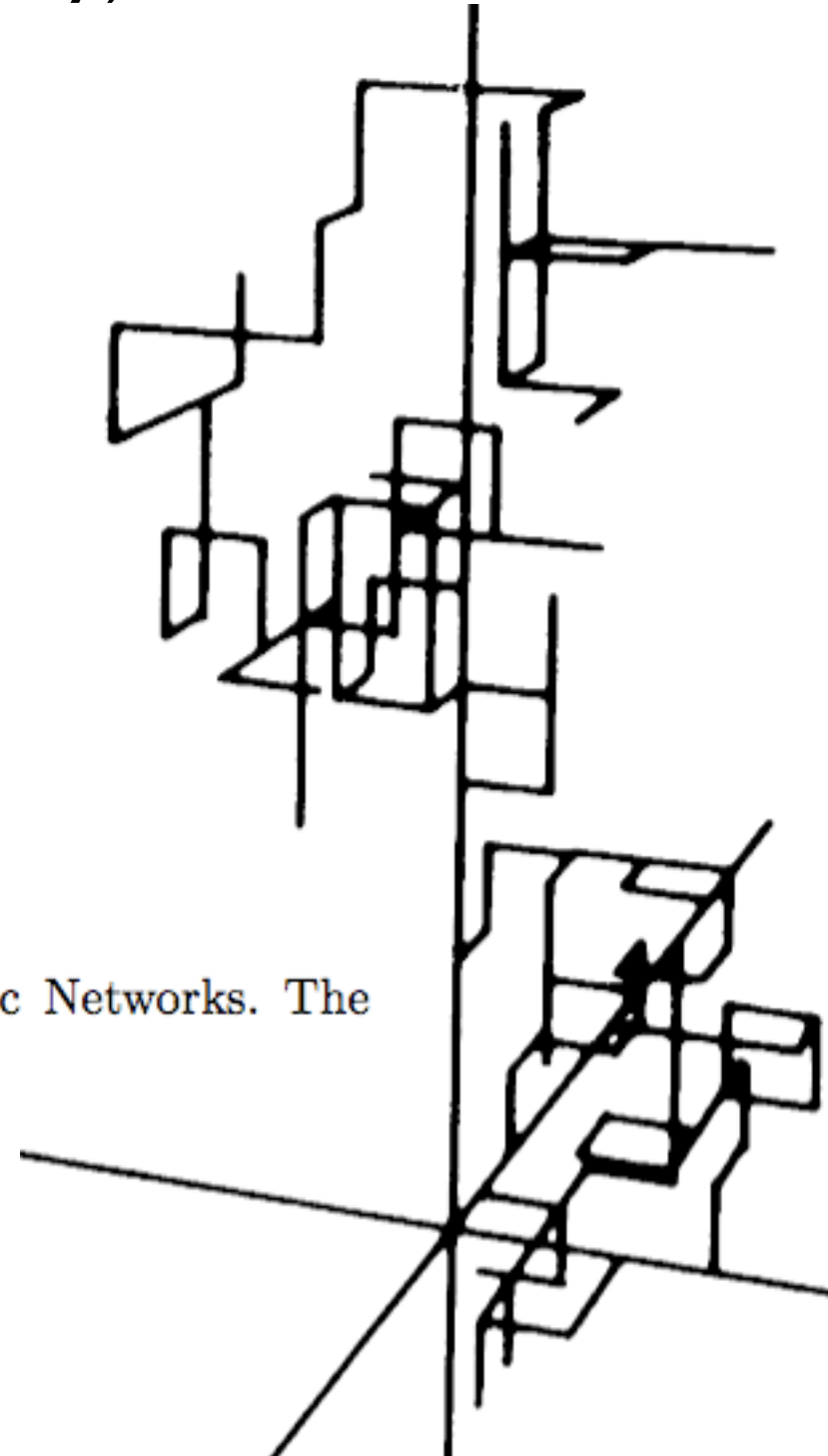
Ether Dipsomania is a term related to an uncontrollable craving for a consistent and appealing theory of something in the form of an analogon to the theory of electromagnetism...the obsession is so compulsive that the ether dipsomaniac will....

Infinite networks are (mathematically) more fun!

Poyla's Theorem:

The Escape Probability is
0 in 2 dimensions, but
0.65946267046 in 3 dimensions

P.G. Doyle and J. L. Snell. Random Walks and Electric Networks. The Mathematical Association of America, 1984.



Appendix

The second fundamental mystery of probability theory is that:

$$E(X) = E(E(X | Y)).$$

As an application let

$$X = \sum_{k=1}^N R_k$$

where N is a positive integer valued random variable and let the $\{R_i\}$ share the expected value $E(R)$. Then

$$E\left(\sum_{k=1}^N R_k\right) = E\left(E\left(\sum_{k=1}^N R_k \mid N = n\right)\right)$$

2FMPT

$$= \sum_{n=1}^{\infty} E\left(\sum_{k=1}^n R_k\right) P(N = n) = \sum_{n=1}^{\infty} n E(R) P(N = n)$$

1FMPT

$$= E(R) \sum_{n=1}^{\infty} n P(N = n) = E(R) E(N)$$

Wald's Theorem (Renewal Theory)