

デジタルドキュメントネットワーク構築システムにおける

力学モデルを利用した表示の試み

Visualizing Digital Document Network by Using Force-directed Placement

松永 賢次

Kenji MATSUNAGA

専修大学 ネットワーク情報学部

School of Network and Information, Senshu University

要旨:

我々は、多くの資料がデジタルドキュメントとして個人のストレージに多く保有されている状況に着目し、それらのドキュメントにメモをつけ、複数のドキュメントに存在するメモ間にリンクを付与することでネットワーク化する「デジタルドキュメントネットワーク構築システム」を研究してきた。本稿では、そのシステムのために、作成されたデジタルドキュメントネットワークを、「力学モデルによるグラフ描画」のライブラリ(d3.js)を利用し可視化した試みについて報告する。辺が交差してしまう複雑な関係を持つ、プログラミング科目のドキュメントネットワークを例にとり、二次元上でわかりやすく表示できたことを示した。

Abstract:

With the increasing volume of privately-owned digital documents it is becoming more important to manage the documents for effective use. For document management, we proposed a digital document network system. In this paper the network viewing functions added to the system are explained. By visualizing with “Graph Drawing by Force-directed Placement,” it is possible to display a complicated digital document network in an easy-to-understand manner and to recognize target nodes interactively.

1. はじめに

近年、多くの資料がデジタルドキュメントとして個人のストレージに保有されており、今後もデジタル化された文書や資料の量は、増加の一途をたどると予想される。従来から紙媒体の書籍や資料は、個人の知識を整理するための有用な知識源であり、人類は紙ベースの書籍や資料を利用して知的作業を行ってきた。今後、紙の文書に代わる大量のデジタルドキュメントをもとに、知識を整理したり新しい意味を加えたりする知的作業をすることが求められている。そのためには、従来の作業方法をもとにデジタルの利点を活用し支援するシステムが必要である。そこで我々は、個人の知識を整理するための「デジタルドキュメントネットワーク構築システム」を研究してきた[1][2][3][4][5]。

これまで開発してきたシステムの編集アプリケーションを利用して、デジタルドキュメント間を結ぶネットワークを構築することができるものの、ドキュメント間の関連が多くなるにつれ、どのようなネットワークを構築したのか、ユーザがイメージすることが難しくなっているという問題（ハイパー空間の迷子問題）が起こっていた。

辺が交差する可能性があるネットワークグラフを二次元表示するアルゴリズムとして「力学モデルによるグラフ描画」がある。このアルゴリズムを実装した JavaScript のライブラリ(d3.js)を利用し、これまで作成してきたデジタルドキュメントネットワークを可視化することを試みた。

2節では、我々が研究してきたデジタルドキュメントネットワークシステムの概要と問題点について述べる。3節ではネットワークの可視化に求められる要求事項について述べる。4節では「力学モデルによるグラフ描画」を紹介し、5節では、辺が交差してしまう複雑なプログラミング授業用のドキュメントネットワークを例にとり、d3.jsのライブラリを活

用して、どのように二次元表示できたか示す。最後に6節では、まとめと今後の課題について述べる。

2. デジタルドキュメントネットワーク構築システムの機能と問題点

我々が研究しているシステムでは、用意したデジタルドキュメントに対して、「マーク」と「リンク」を利用者がつけていく。マークを付加することは、紙の資料に付箋を付けたり、下線を引く行為に相当する。本システムの編集アプリケーションを利用するユーザは、文書内のある対象に付けるマークを生成し、そのマークにはメモを記述できる。二つのマーク間に、リンクを付加することで、それらの間に関係があることを示すことができる。リンクには、マーク同様メモを付与できる。

リンクは、デジタルドキュメントをまたぐマーク間でも付与することができるので、マークとリンクを付与していった複数のデジタルドキュメントはネットワークを形成することになる(図1)。このようなネットワークを形成する行為は、知識を構築していると見なすことができる。ユーザは、構築されたネットワークを対象として、マークを検索したり、あるマークを起点としてリンクをたどりながら別のマークに移動したりすることで、複数のドキュメントまたぐ知識を呼び起こすことができる。

以上述べたことをまとめると、本システムは利用者から見たときに8つの基本機能を持っていると言える[4]。

1. 文書を登録する。
2. 文書を見る。
3. 文書にマークをつけ(生成し)、マークに対してメモを記述する。
4. 画面を見ながらマークを選ぶ。

5. マークを検索する。
6. 二つのマークの間にリンクをはり、リンクに対してメモを記述する。
7. リンクをたどってマークを選ぶ。
8. リンクを検索する。

本システムを利用して、デジタルドキュメントネットワークを構築することは、知識を整理する学習活動と見なすことができ、個人のみならずグループ学習にも有益である[4][5]。学習者たちは、ドキュメントを用意し、メモをつけ、リンクをはるという行為を相談しながら行っていけば、より深く知識を整理することができる。

本システムの画面は、図2に示すように、ドキュメントの

ページを表示したときにマークの一覧が出たり、リンクの一覧が表示されるものとなっており、規模が大きくなるにつれ、作成しているネットワークがどのようなものとなっているのか、関係性を図1のような形でイメージすることが難しくなっているという問題がある。

ネットワークを記録できるということが、本システムの利点であるものの、記録したネットワークを、必要に応じてユーザ（学習者）が想起できなければ、整理した知識を活用できないということになる。そこで、本システムにおいて、ネットワーク空間をユーザが把握しやすいうように表示する機能が必要となっていた。

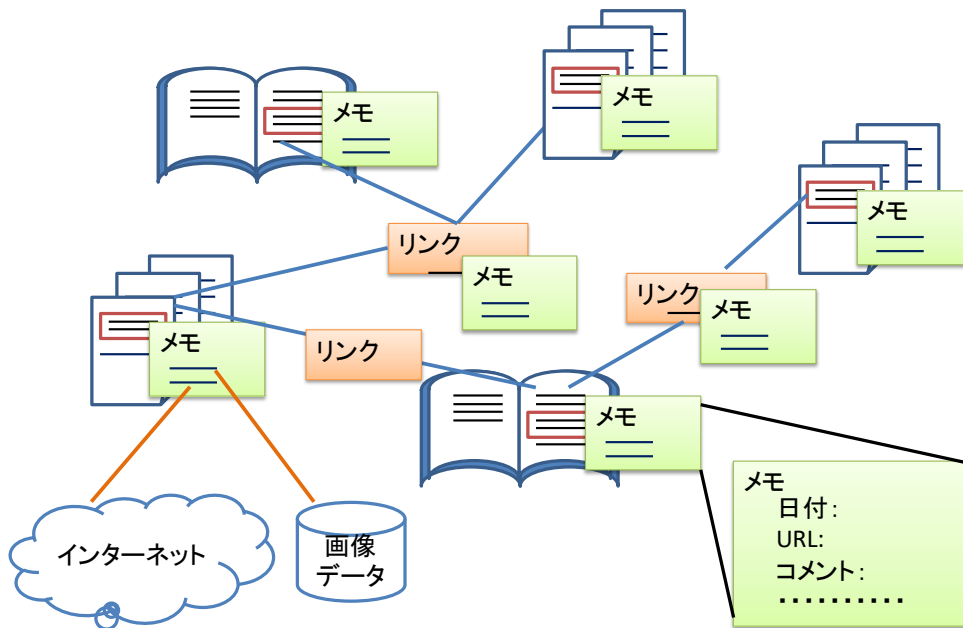


図1 デジタルドキュメントネットワークの概念

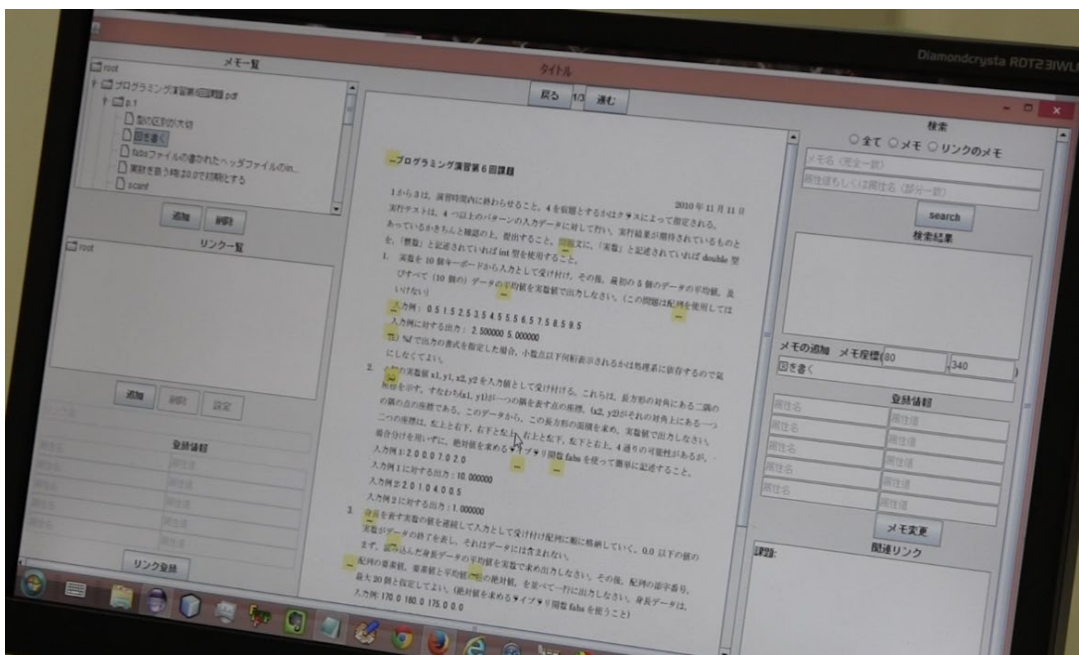


図2 デジタルドキュメントネットワーク構築システムの編集画面

3. デジタルドキュメントネットワークの表示への要求

本システムを用いて作成する「デジタルドキュメントネットワーク」の性質について述べ、そこから、ネットワーク表示機能に求められる要件について検討する。

デジタルドキュメントネットワーク構築システムでは、図3左に示したように、ドキュメント内にあるメモ同士にリンクをはるようになっていて、メモは、ドキュメント内に存在するため、ドキュメントとメモの間には包含関係があり、その包含関係もつながりとして表現できる。そのように考えると、図3右にあるような形で表現できる。文書の数が増えると、ネットワークの形状は、グラフ理論で言うところの一般的なグラフとなり、木構造のような単純なものとなっていない可能性がある。木構造となっていないネットワークを二次元平面に表示した場合、辺（リンク）同士が交差してしまう可能性がある。

ユーザにより、リンクは随時、追加されていくものである。リンクが追加されるたびに、ネットワーク形状の表示が適切に変わりうるアルゴリズムを用いなければならない。

個人または少人数のグループによって構築していく閉じたネットワークであるため、ネットワークの大きさ（超点の数、辺の数）は、ある程度の範囲で収まると考えてよい。そのため、使用するアルゴリズムの計算量に対する制約は厳しくはない。着目しているドキュメントの周辺のみ見れば良いと考えれば、ノード数を制限して表示することもできる。

4. 力学モデルによるグラフ描画

動的にネットワーク表示を構築する代表的な手法として、力学モデルによるグラフ描画(Force-directed Graph Drawing)がある。この方法は、無向グラフを描くときに次のことを目標としたものである[6]。

1. 頂点（ノード）を、フレーム内に等しく分布させる。
2. 辺（エッジ、リンク）の交差を最小限にする。
3. 辺の長さを均一にする。
4. 内在している対称性を反映させる。
5. フレームにあうようにする。

力学モデルによるグラフ描画のアルゴリズムはいくつか提案されているが、辺をバネに見立てるような物理的なアナロジーが使用され、力学的に安定した配置になるまでシミュレーションを行うといったものとなっている。

シミュレーションを行って配置を決定していくアルゴリズムであるため、利用者がマウス等を使用し、対話的に頂点の位置を変更すれば、配置を再計算させることで、その挙動を観察するといったインタラクティブな使用方法もできる。

図4に、力学モデルによるグラフ描画によりネットワークを表示するAPI（ライブラリ）の代表的なものである、D3.jsというJavaScript言語用のライブラリ(API) D3.js[7][8]を利用して、図3右の図を描いてみたものを3パターン示した。ノードの初期位置をランダムに決めて計算するため、異なる配置となる。インタラクティブにノードを動かして再計算させても異なる配置となる。

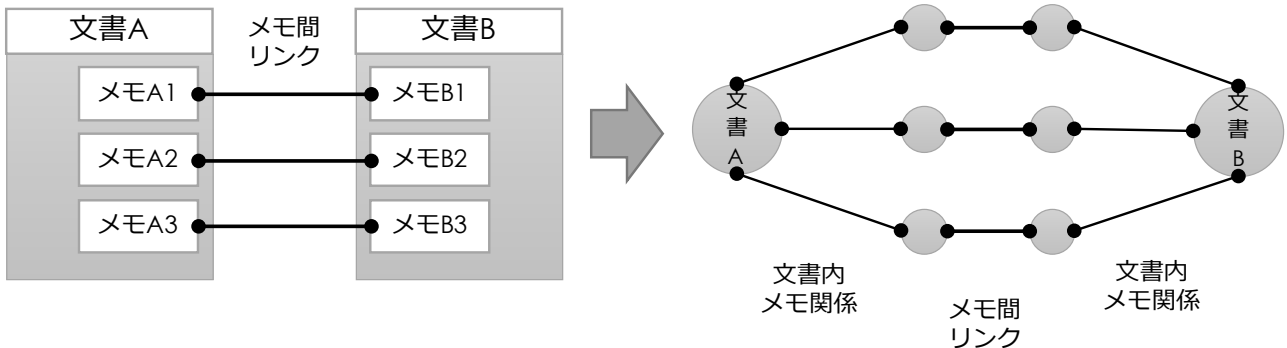


図3 デジタルドキュメントネットワーク構築システムの内部要素の関係とそのグラフ表現

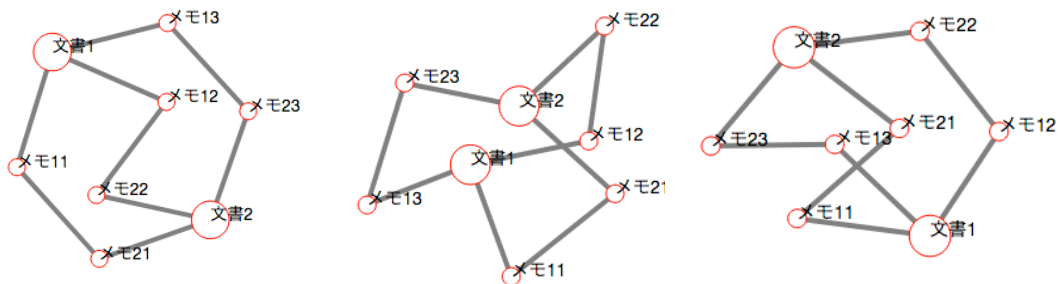


図4 d3.jsの力学モデルによるグラフ描画を利用して、図3のグラフ表現を描いた例

5. プログラミング科目ドキュメントを対象とした表示

5.1. 対象事例の説明

力学モデルによるグラフ描画は、再計算により同じ形はならないという欠点はあるものの、D3.jsを使用することで、本システムで構築したネットワークの表示に実際に適しているか確認する実験を容易に行えるという利点がある。また、計算量も実質的には問題とならない。

本節では、専修大学ネットワーク情報学部のパログラミングの導入科目で使用したドキュメントに対して、学生たちの少人数のグループによってメモとリンクが作られたネットワークを使用した例を用いたものを示していく。このネットワークを取り上げるのは、比較的小み入ったリンク関係となっており、見やすく表示することが比較的困難と考えたからである。

ドキュメントが使用されたプログラミング科目は、週に2回の授業から構成されており、1回の授業が座学教室での講義スタイル、もう1回の授業が端末室での演習スタイルで実施された。講義授業では、スライドをPDFにしたドキュメントが与えられる(図5左)。演習授業では、作成するプログラムの仕様が記述された問題が、ドキュメントとして与えられる。

これらの授業間では、典型的に次の関係がある(図5右)。

- 講義と演習との関係: 演習の問題文では、講義で示された用語・概念で記述される。問題内容が、講義で示され

た例題を活用したものになっている。

- 前後の講義間での関係: 後の講義での記述は、前の講義で定義された用語で記述される。前の講義の例題を発展させたものが、後の講義の例題で示される。
- 前後の演習間での関係: 後の演習での問題では、前の演習の内容を発展させた問題が出題されたり、前の演習の問題を複数融合した問題が出題される。

学生たちは、以上に示したような関係を、ドキュメントネットワークシステムの編集アプリケーションを利用して、メモとリンクで記述している。

5.2. D3.jsを使用した力学モデルによる表示

ネットワーク表示のプロトタイププログラムを、[8]で示されているコードを元に改良して作成した。メモとリンクの記述データをjson形式に変換したものを入力とし、配置に関するパラメータを、APIのデフォルト値そのまま使用したところ、中央部に固まってしまう見にくくなってしまった。

そこで、試行錯誤的にいくつかのパラメータを調節し、このドキュメントネットワークグラフを人間が見て理解可能な状況になるように試みた。辺の長さ及び電荷値(charge)のパラメータを調整することで、図6のようにある程度見やすいものとなった。電荷値の値は、負の方向に大きくすることで、頂点同士が反発して離れることになる。一方、あまり大きくすると、画面に広がることになってしまうため、見やすさと、表示するグラフ範囲のコンパクトさとのトレードオフとな

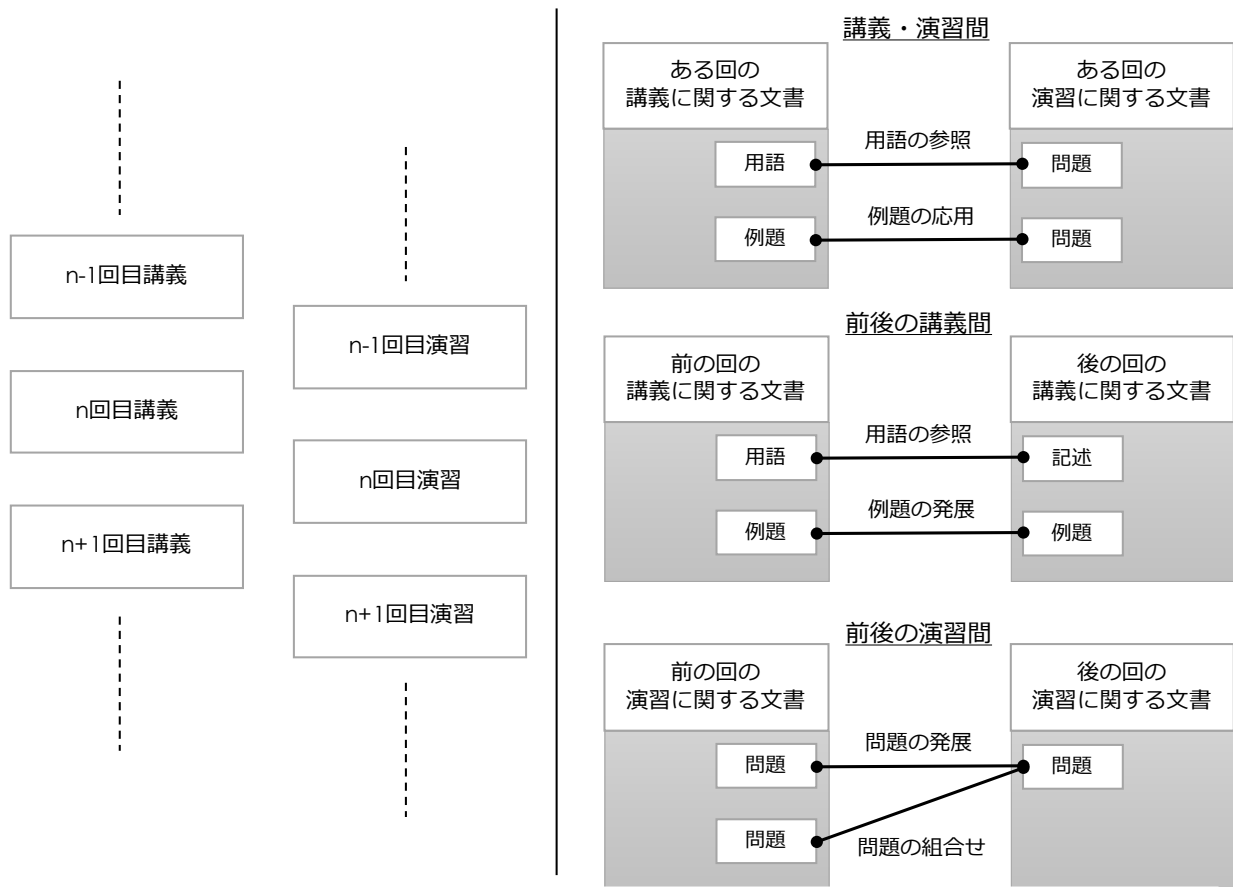


図5 プログラミング科目のドキュメントの関係と代表的なメモとリンク

っている。

図6のように、ドキュメント名やメモのタイトルを表示すると、概観するのは適当だが、重なって読みづらいという問題点がある。そこで、[9]で示されているコードを参考に、マウスオーバーすることで文字列の表示をし、文字列検索によるノードの検索できるように改良を試みた(図7,8)。

辺の長さを固定とせず、必要に応じて長くすることができるようにしていることも、辺の交差が少なくなり、わかりやすいレイアウトとなる要因になっている。ドキュメント(大きな円)同士においても、連続した講義と演習のドキュメントが近い配置になっており、よりわかりやすくなっている。

図8は、「文字列」を含むメモ(小さい円)を検索している様子を示している(わかりやすいように画面の一部を切り取っている)。検索されたメモは、円の色が変わって表示される。その円のところにカーソルを載せると、メモに記載されている文字列を見ることができるようになっている。これにより、ある文字列を持ったメモが、どのドキュメント周辺に多いのかわかりやすくなり、ドキュメント間の関係を理解する助けになっている。

6. 最後に

力学モデルにより、構築されたデジタルドキュメントネットワークの概観を理解しやすくなることを、複雑な関係を持っているプログラミング科目のネットワークを利用して示した。

今後は、この表示機能を編集機能と統合するとともに、より大きなネットワークを対話的に調べていけるようにしていきたい。

参考文献

[1] Yoshikoshi, M., Matsunaga, K., Yoshida, K. “A personal document network building system for digital document

searches,” *HCI 2013, Part II. CCIS*, Vol.374, pp.458-461, 2013.

[2] 松永賢次, 吉越将紘, 吉田享子 “個人の知識を整理するためのデジタルドキュメントネットワーク構築システムの提案,” 情報処理学会 研究報告情報システムと社会環境 (IS), Vol. 2013-IS-125, No.10, pp.1-5, 2013.

[3] Matsunaga, K., Yoshida, K. “Digital document network system for organizing individual knowledge,” *HCI 2014, Part I. LNCS*, Vol.8521, pp.396-403, 2014.

[4] 松永賢次, 吉田享子 “グループ利用のためのデジタルドキュメントネットワーク構築システムの機能と実装について,” 情報科学研究所報, No.85, pp.9-15, 2015.

[5] Matsunaga, K., Yoshida, K. “Using digital document network system for group learning activities,” *HCI 2015, Part II. CCIS*, Vol.529, pp.1-5, 2015.

[6] Fruchterman, Thomas M. J., Reingold Edward M. “Graph Drawing by Force-directed Placement,” *Software-Practice and Experience*, Vol.21, No.11, pp.1129-1164, 1991.

[7] D3.js Data-Driven Documents, <https://d3.js.org>

[8] 森藤大地, あんちべ, エンジニアのための データ可視化[実践]入門 ~D3.js による Web の可視化, 技術評論社, 2014.

[9] Vallandingham, J. “How to Make an Interactive Network Visualization,” <https://flowingdata.com/2012/08/02/how-to-make-an-interactive-network-visualization/>, 2012.

付記

本稿は、平成 28 年度専修大学研究助成「デジタルドキュメントネットワークの視覚化と検索手法に関する研究」による成果の一部である。



図6 プログラミング科目のドキュメントネットワークを d3.js の力学モデルライブラリで表示した例

ドキュメントネットワーク

Document Network

プログラミング

Search

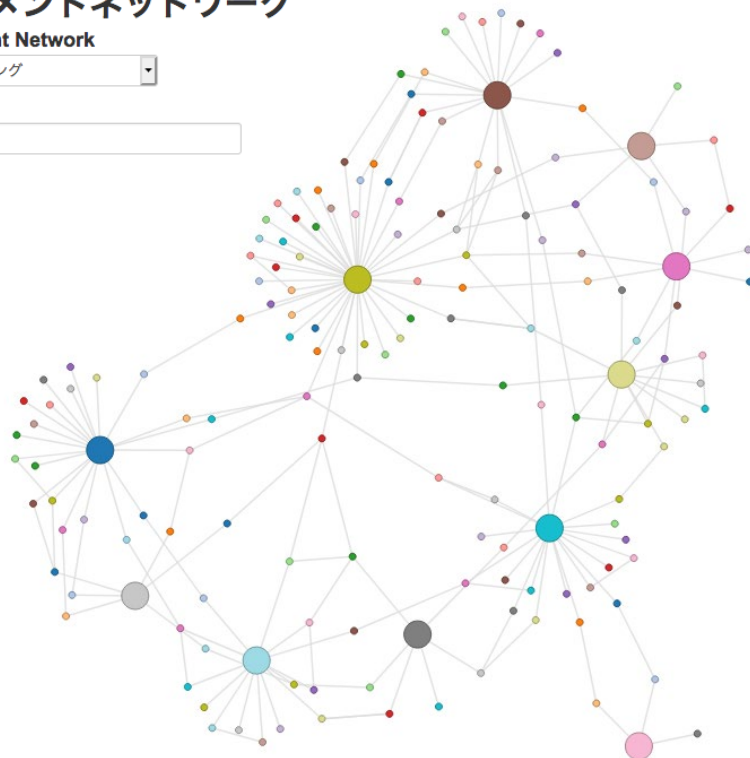


図7 図6で示したネットワークの表示を改良したもの

ドキュメントネットワーク

Document Network

プログラミング

Search

文字列

文字列

文字列の比較
142 in プログラミング入門2010-10.pdf(11)-P.6



図8 文字列検索によって図7のネットワーク内を調べる例