

複数医療機関の頻出医療行為の利便性を向上した視認性の高い可視化手法

澤村 今日子 (指導教員：文化情報工学科・Le Hieu Hanh)

1 はじめに

近年、医療行為の効率性・安全性向上のために、単一の医療機関の電子カルテデータに分析が進んでいる [1]。また、特定疾病に対する医療行為の比較のために、複数医療機関の電子カルテデータに分析も行われている。安光らの研究では、複数医療機関の電子カルテデータにシーケンシャルパターンマイニングを適用し、それぞれの医療機関のシーケンスバリエーション (以下 SV) を抽出し、それらに対してクラスタリングを行い、同じグループに所属された SV に対して生成された併合シーケンシャルバリエーション (以下 MSV) のテキストファイルが算出された [2]。しかし、この研究ではこの MSV の可視化がまだ手動で行われており、人的コストも上がり、ミスの危険性も高まるという問題点がある。これまで、他に可視化 web アプリケーション [3] が提案されたが、データ形式が MSV のデータと互換性がなく、加えて複雑な分岐に対応していないためそのまま適用できない。

そこで、本研究は MSV の自動可視化 web アプリケーションを提案する。本提案により、より手軽に医療従事者が他医療機関の医療行為を比較できるため、医療行為のミス削減が期待できる。また、とある疾病に対する国内の一般的な医療行為が一目でわかるため、まだ歴の浅い新人医療従事者の勉強にも使用できる。

2 提案手法

本節は提案手法について説明する。まず、本提案手法の想定する利用場面のついて説明したあとに、可視化の流れを述べる。次に、可視化した MSV の例を紹介し、開発した可視化ツールの特徴を説明する。

2.1 想定する利用場面

医療従事者はパソコンやタブレットなどの電子機器を使用して、本可視化アプリケーションを使用する。医療従事者は同じグループに分類された複数医療機関のある疾患の治療に関する MSV を確認し、医療機関間の共通医療行為と差異医療行為を簡単に確認することができる。それによって、対象となる疾病の特徴を瞬時に把握することができ、医療改善につながる。

2.2 可視化の流れ

図 1 は可視化手法的な流れを表示する。MSV テキストファイルを自動的に可視化しやすくするために、json ファイルに変換する。そして変更した json ファイルを Javascript を用いて可視化する。今回は web アプリケーションの開発であるため、最終的に HTML 形式のファイルで出力される。

2.3 json 変換

MSV テキストファイルの一例の内容を図 2 に示す。このファイルには複数のノードの情報を含む。その中で、一つノードには name, label, id, nextNodes の 4 つの要素が記載されている。name はノードの名前と日にちの情報からなる。label はこのノードがどの



図 1: 可視化の流れ

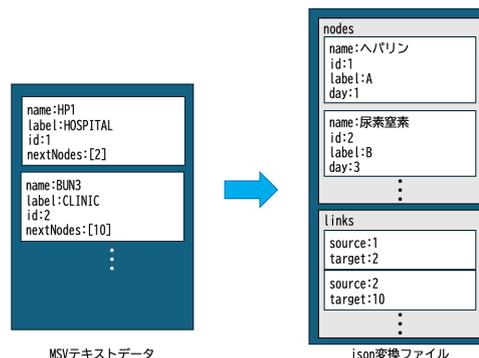


図 2: json 変換

SV に所属しているかを示す。id がそのノードの独自の識別子であり、nextNodes はこのノードと繋がり先の複数ノードの id を含む。この情報を用い json ファイルに変換する。変換後の json ファイルにはノード群の情報を管理する nodes とノードとノードの間の枝を管理する links の 2 種類の情報を生成する。nodes には各ノードの情報は name, id, label と day が含まれており、links にはノードとノードの間の枝の全ての情報が作成される。

図 2 について説明する。左が元の txt データで、右が変換した json データである。1 番上のデータを例とする。nodes には上記の 4 種のデータが含まれる。name は、元は英語の略称名「HP」と格納されている。それを json データでは日本語名「ヘパリン」に変換して格納する。id は 1 から順に並んでいない場合は数字を新たに割り振り直し格納する。元のデータでは「HOSPITAL」となっている label であるが、json データでは匿名処理を行ってから label に格納する。この作業により医療データの機密性が守られる。また、name の末尾の数字は day 情報を示している。そのためこの例では day が 1 となる。そして links には source と target の 2 種の情報が含まれる。source には自身の id 情報が、target には nextNodes 情報がそれぞれ格納される。

2.4 可視化処理

第 2.3 節で生成された json ファイルから d3-sankey ダイアグラム [4] の形で自動的にグラフ化を行う。ノードが並べられるアルゴリズムとしては、まず start と end を両端に設置する。そして次に day の値が小さい順に左から並べていく。また、SV ごとに高さを設定



図 3: 可視化結果

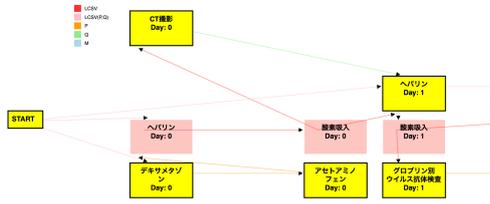


図 4: ハイライトされたノード

する。

ここでは、視覚的にわかりやすくするために、label の情報を利用しノードと枝の色を決める。同じ SV に所属するノードと枝が原則的に同じ色にする。さらに、視認性を高めるためにノードの上に医療行為の略語ではなく、下の医療行為名前を表示するように工夫する。これによって、誰でもグラフを見るだけでノードの意味をすぐ把握することができる。

また、ノードをクリックするとそのノードと同じ医療機関のノードがハイライトされる。例えば医療機関 A をクリックすると、他の全ての医療機関 A のノードと、医療機関 A が含まれた、全ての最長共通サブシーケンスバリエーション (以下 LCSV) のノードがハイライトされる。

3 実験

本節では提案手法の有効性を確認するために、実データを用いた評価実験を行う。

3.1 データセット

今回の実験では、「COVID-19 感染の診療への影響調査と予測モデル開発」に関する研究の研究協力機関である 23 医療機関 (医療機関 A, B, ..., W) から提供された COVID-19 の第 5 波 (2021.07.01-2021.09.30) に対する実電子カルテデータ中の医療指示データを対象とした。その中で、同じグループに分類された医療機関 M, P, Q の MSV の結果を自動的に可視化することを目指した。本研究は一般社団法人ライフデータイニシアティブの利用目的等審査委員会による審査を受け、承認された上で研究を進めた (審査番号 2021_MIL_0007_A002)。

3.2 実験結果

図 3 に示した通り、医療機関 M, P, Q の MSV を自動的に可視化した結果が表示された。3 つの医療機関の共通 SV は、赤色で表示される。そして、医療機関 P と Q の共通 SV に属するノードはピンク色で区別することができる。その他、医療機関 M, P, Q のみの SV にある医療行為はそれぞれ水色、オレンジ色、薄緑色で描かれている。色に加えて、各ノードには医療行為の名前と発生した日にち情報も記載されているためグラフを見る人は誰でも SV の内容を簡単に把握

することができる。また、ノードをクリックするとそのノードと同じ医療機関の情報が含まれているノードがハイライトされた。例えば医療機関 P をクリックすると、医療機関 P のノード全てと、LCSV(P, Q) と、LCSV(P, Q, M) のノードがハイライトされる。このように、MSV テキストデータから手動でグラフ化を行った既存研究と比べて、高い視認性のグラフを自動的に生成することができることを確認できた。

4 おわりに

本研究では、複数医療機関間の MSV を自動で可視化処理する web アプリケーションを開発した。手法の流れとしては、MSV のテキストファイルに含まれる医療行為情報、日付情報、ノード情報を、d3-sankey でプロットできる形に json ファイルとして F 変換した。その json ファイルを Javascript を用いて d3-sankey の形で可視化処理を行い、html ファイル形式で出力する。これによって、MSV の可視化も早く、正確に、わかりやすくなり医療界に対して大きく貢献できると確認できた。

可視化グラフにはより高度なインタラクティブな機能を開発する予定である。例えば、ユーザはある部分のみ着目したい場合に、ノードをクリックすることでより詳細な情報を提示させたりできるような機能も考えられる。また、実際に医療従事者に使用していただき評価をいただくことにより医療現場で使用しやすいものになるように開発する。

謝辞

本研究の一部は日本学術振興会科学研究費 (#24K02943) の助成からの支援によって行われた。

参考文献

- [1] 横田治夫. 電子カルテデータ解析-医療支援のためのエビデンス・ベースド・アプローチ-. 共立出版, 2022.
- [2] 安光夕輝, Le Hieu Hanh, 松尾亮輔, 山崎友義, 荒木賢二, 横田治夫. クラスタリングを用いた多病院間の頻出医療指示パターン比較. In *DEIM Forum*, No. 5b-6-3, 2023.
- [3] 小林梨華, 坂本任駿, Le Hieu Hanh, 荒木賢二, 横田治夫. 動的な患者情報を用いた医療行為推薦を支援するための医療シーケンスの可視化. In *DEIM Forum*, No. C25-4, 2021.
- [4] d3/d3-sankey: Visualized flow between nodes in a directed acyclic network. <https://github.com/d3/d3-sankey>. オンライン, アクセス 2025 年 1 月 8 日.