

# 相関ルールによる唐代官僚遷転\* の分析

白須裕之<sup>†</sup>  
京都大学人文科学研究所

永田知之  
京都大学人文科学研究所

## 概要

データマイニングの主要な要素技術として、相関ルールマイニングの方法がある。本稿はこの方法を使って、中国唐代（618-907）の官僚が遷転した様子を解明しようとする一つの試みである。官職の遷転を研究する上で重要な要因としては、「エリートコース」として考えられる官歴のパターンがある。しかし、史料上、このような「エリートコース」が明示的に記載されることは稀である。人物が任官したという情報をデータベースとすると、そこから任官の傾向を示すルール「官職 A, B についたことがある人はしばしば、官職 C につく」を抽出することができる。このようなルールで抽出された官歴パターン中に、史料に明示されていない有意なパターンが含まれることが考えられる。

## Association Rules for Appointment Transactions of Government Officials in the Tang Dynasty

SHIRASU Hiroyuki\*  
Institute for Research in Humanities  
Kyoto University

NAGATA Tomoyuki  
Institute for Research in Humanities  
Kyoto University

### Abstract

Association rule mining is a popular knowledge discovery technique. We merely discover elite courses of government officials in the Tang dynasty on historical documents. The paper suggests the first step of applications of association rule mining to investigate appointment transactions of their government officials. We are given a database of appointment transactions in which each transaction consists of government officials hold by a person in his life. This paper presents elite courses by mining association rules from the database.

## 1 はじめに

データマイニングの主要な要素技術として、相関ルールマイニングの方法がある。本稿はこの方法を使って、中国唐代（618-907）の官僚が遷転した様子を解明しようとする一つの試みである。官職の遷転を研究する上で重要な要因としては、「エリートコース」として考えられる官歴のパターンがある。人物が任官したという情報をデータベースとすると、そこから任官の傾向を示すルール「官職 A, B についたことがある人はしばしば、官職 C につく」を抽出することができる。このようなルールの形で抽出された官歴パターン中に、史料に明示されていない有意なパターンが含まれるこ

とが考えられる。

以下は本稿の構成である。節 2 では唐代官僚の遷転についての背景を述べ、官歴を研究する上で重要となる要因を提出する。節 3 では、唐代官僚の遷転研究における先駆的な業績である文献 [10] 『唐代中央重要文官遷転途徑研究』での官歴の扱い方、及び同書の問題点を明らかにする。また「官歴のパターン」として時系列の情報を扱うことができない理由を述べる。節 4 では相関ルールマイニングの概要を本稿に必要な範囲で紹介する。節 5 では相関ルールを使って、唐代官僚の遷転の様子をいかに分析したかについて述べ、節 6 で本稿の結論と課題を述べる。

## 2 唐代官僚の遷転について

唐代の制度においては、現代日本でいうところのキャリア官僚の就く官職を正一品から従九品下の 30 階級にランク付けしていた。このランクを以下、品階と呼

\*本稿では官僚の官職間における着任（‘任’、‘拜’、‘授’、‘叙’等）・昇任（‘擢’等）・横滑り（‘遷’、‘転’等）・降格（‘貶’、‘降’、‘左遷’、‘左授’等）・離職（‘辞’等）といった異動をまとめて‘遷転’と称する。

<sup>†</sup>京都大学 21 世紀 COE 東アジア世界の人文情報学研究教育據點

\*Toward an Overall Inheritance and Development of Kanji Culture

ぶ。官職を「一品」を最上位として、「九品」までの九つの階級に分け、また、各「品」を「正」を上位として、「正」「従」の二つの階級に分ける。これを「正一品」、「従一品」などと呼ぶ。さらに「正四品」から「従九品」について「上」を上位として、「上」「下」の二つの階級に分ける。これを「正四品上」、「正四品下」などと呼ぶ。こうして得られた官職の階級が品階である<sup>1</sup>。

それでは、品階が高い官職に就任しさえすれば、官界での栄達を望む官僚たちにとって、その自尊心は満たされるのかといえば、事情はそう単純ではない。例えば、8世紀前半に著された『御史台記』<sup>2</sup>という文献には、次のような一節が見られる。

世間で取り沙汰するところでは畿県尉から次に遷る官職には六つの（お決まりの）コースがある。監察御史になるのは「仏道」、大理評事になるのは「仙道」、京県尉になるのは「人道」、畿県丞になるのは「苦海道」、県令になるのは「畜生道」、諸州諸曹参軍になるのは「餓鬼道」という。<sup>3</sup>

ここに挙げられる畿県尉は正九品下、そこから従八品下の大理評事へ遷るのが「仙道」でありながら、同じ従八品下の京県尉へ遷るのが「人道」、大理評事や京県尉より品階が二つ上である正八品下の畿県丞へ遷ることが「苦海道」と称されている。ここからも品階の上下だけが、官僚たちの考える官職の良し悪しを示す指標ではなかったことが分かる。

当時、「清望官」、「清官」と称される官職が存在した。これらは現代の言葉でいえば、「エリート」に相当する。品階が低くとも「清官」である官職は存在したし、また比較的それが高くとも「清官」とは認識されない場合は往々にしてあり得た<sup>4</sup>。

これに加えて、部署の問題も見逃せまい。「前行要望」と呼ばれる花形部署への配属が、官僚社会で歓迎を受けていたことは想像に難くない。一方でそれ以外の部署が「後行閑司」と称され、職務がさほど繁忙ではないと意識されていた事実も、今日に伝わっている<sup>5</sup>。

さて、今まで述べてきたのは個々の官職に「エリート」、「非エリート」（これはもちろん両者を比較した上

での定義であり、官僚である以上、当時の社会においては大変な選良であったことは間違い無い）の区別が存在したということである。それでは官職の間における遷転の経路に優劣は無かったのか、いわゆる「エリートコース」はそこに見られたのであろうか。実はこれについても、いくつかの記録が残っている。

- A 1 進士（任官試験である科挙進士科の及第者）→ 2 校書（所属する部署によって品階は異なる。崇文館校書：従九品下、弘文館校書：従九品上、秘書省校書郎：正九品上）→ 3 畿県尉（正九品下）→ 4 監察御史（正八品上）→ 5 拾遺（従八品上）→ 6 員外郎（従六品上）→ 7 中書舍人（正五品上）→ 8 中書侍郎（正四品上）
- B 1 制策（任官試験である科挙制科の及第者）→ 2 正字（所属する部署によって品階は異なる。太子司經局正字：従九品下、秘書省正字：正九品下）→ 3 畿県丞（正八品下。一説には京県尉：従八品下）→ 4 殿中侍御史（従七品上）→ 5 補闕（従七品上）→ 6 郎中（従五品上）→ 7 給事中（正五品上）→ 8 中書令（正三品）

これは唐代の文献に見られる宰相（皇帝を補佐し、官僚機構の頂点に立つ。唐代では常時、複数名が任ぜられた）にまで至る人物の典型的な昇進の経路である。Aには「八儁」という独自の呼称すら存在しており、Bと合わせて当時の人々から「エリートコース」と考えられていたことは確かであろう<sup>6</sup>。

もとより、宰相になる人物が全て「八儁」の経路に沿って、昇進したわけでもないだろう。しかし、開元十年（722）頃に成立したとされる地理書『兩京新記』<sup>7</sup>には、「官職異動の経路で、拾遺・大理評事・京県尉の三つの職を経るのは、時の人から栄えあることとされた」<sup>8</sup>という。将来における出世のために、経験しておくべきだと意識される官職は、やはり存在したのである。

「エリートコース」などというものは、いつの時代、どの社会においても厳然として存在するのかもしれない。ただし、それらが制度として記録、それも公式編纂物に見られることは、ごくまれである。全てのポストに存在意義を認める建前の下、組織自らがある職務のみに特別な意味を付与することは許されないからだ。そ

<sup>6</sup> 文献 [13] 頁 16、17 を参照。

<sup>7</sup> 『兩京新記』については文献 [12] を参照。

<sup>8</sup> 『兩京新記』は既に散逸。ここに挙げた一節は南宋の謝枋得が編纂した『秘笈新書』に引用されて伝わっている。文献 [8] 頁 211、269、436 を参照。

<sup>1</sup> 唐代の官僚制度に関する研究は少なからずあるが、本稿は中でも文献 [4]、[11] に負うところが極めて大きい。

<sup>2</sup> 『御史台記』については文献 [3] を参照。

<sup>3</sup> 『御史台記』は既に散逸。ここに挙げた一節は文献 [9] 頁 452、453 に引用される形で残っている。原文「唐畿尉有六道、入御史為仏道、入評事為仙道、入赤尉為人道、入畿丞為苦海道、入県令為畜生道、入判司為餓鬼道。」

<sup>4</sup> 文献 [16] 頁 33、34 を参照。

<sup>5</sup> 文献 [16] 頁 36 を参照。

れだけに‘八儒’にまつわる史料などは極めて貴重なのだが、当時の官界における遷転の経路を究明するにあたって、これら少数の記録だけでは、なおもどかしさは免れない。

以上、唐代官僚の遷転を研究するにあたっての重要な要因について述べた。まとめると以下がその要因である。

- 唐代の官僚制度に明文化されている品階
- 「清望官」、「清官」と称されるエリート職か否か
- 花形部署か閑職部署か
- エリートコース

これらの内、‘エリートコース’については史料上、明示的に記載されていることは稀である。しかも、本節で挙げたわずかに残るその記録も全て8世紀のものであり、時代が偏っている憾みがある。本稿は唐代の人物の官歴事例にもとづいて、その中から‘エリートコース’と呼ばれうるような有意の「官歴のパターン」があるか否か、あればどのようなものかを明らかにする試みであると言える。

### 3 『唐代中央重要文官遷転途径研究』について

文献 [10] 『唐代中央重要文官遷転途径研究』は唐代官僚の遷転研究における先駆的な業績である。本節では同書での官歴の扱い方、すなわちそれを表現した「官職遷転表」について取り上げ、また同書の問題点を明らかにする。これらの問題点は官職の遷転の傾向を分析する上で、我々に課せられた課題である。本稿では「官歴のパターン」として時系列の情報を扱わない。この点について本節がその理由を明らかにしている。

#### 3.1 官職遷転表について

文献 [10] 『唐代中央重要文官遷転途径研究』においても官職の遷転についてのデータをまとめたものが「官職遷転表」である。その形式は以下のようなものである。「官職 X」について同書は「官職 X 遷入表」、「官職 X 遷出表」という二つの図表を作成する。前者では官職 X についての人物及び着任の時期を縦軸に、その人物が官職 X に着任する前に就いていた官職を横軸に配列

する。後者では官職 X についての人物及びそこから離任した時期を縦軸に、その人物が官職 X から遷った官職を横軸に配列する。両者を合わせることによって、どのような官職から官職 X に就任し、また官職 X よりいかなる官職に遷るケースがあったかが示されるようになっている。

表 1: 官職遷転表の例

人名	時期	官職名		
		官職 A	官職 B	官職 C
P	T1	✓		
Q	T2		✓	
R	T3	✓		✓

表 1 は同書の「官職遷転表」の例である。表 1 を「官職 X 遷入表」として読む場合、以下のようなことが分かる。人物 P, Q, R の官職 X に着任した時期は各々 T1, T2, T3 である。また、そのときの人物 P が任官していた官職は A、人物 Q は官職 B、人物 R は官職 A, C である。但し、官職 X の着任以後に各々の官職を退任しているかどうかは分からない。

同様に表 1 を「官職 X 遷出表」として読む場合、以下のようなことが分かる。人物 P, Q, R の官職 X を退任した時期は各々 T1, T2, T3 である。また、そのときの人物 P が任官していた官職は A、人物 Q は官職 B、人物 R は官職 A, C である。但し、官職 X の退任以前に各々の官職に着任していたかどうか、またこの時期に着任したのかどうかは分からない。

表 1 の例に出現する官職 A, B, C と官職 X は必ずしも時系列を成す訳ではない。歴史史料の性質上、順序関係が明記されている場合以外は時系列として扱うことはできない。官職遷転表をこのように解釈する理由については以下で説明する。

#### 3.2 問題点

文献 [10] 『唐代中央重要文官遷転途径研究』は今日においても先駆的な業績としての意味を失わない。ただ、刊行より三十年近くを経て、その間、営々と積み重ねられてきた各官職ごとの遷転情報整理の成果が生かされていない<sup>9</sup>こと、使用した史料が根本的なもの<sup>10</sup>に、

<sup>9</sup> 「中央重要文官」に限って言えば、ここにいる「各官職ごとの遷転情報整理の成果」には文献 [2] が挙げられる。

<sup>10</sup> ここで「根本的なもの」とは新田両唐書 [5][17]、及び『唐僕尚丞郎表』 [6] を指す。

対象となる官職が中央の重要な文官にのみ限られていることはその欠点といわなければならない。更にいえば、個人の伝記に「官職 A、官職 B を経験した」とある場合、中国語古典文の性質上、「官職 A から官職 B に遷転した」のか「官職 A と官職 B」を兼任していたのか、この記述だけでは判然としない。文献 [10] はこのような情報に対する解釈に、やや恣意的なところが見られる。また、遷転について仮にその正確な時期情報が割り出し得たとしても、おしなべて当時、在位していた皇帝によってのみ時期は示される（『唐僕尚丞郎表』が詳しく考証してできる限り時期をも示した尚書省の令、左右僕射、左右丞、六部の尚書、侍郎を除く）。これらは多く史料の、物理的な制約によるものであり、同書の情報を増補・訂正するのは、我々後学の者に課せられた務めであろう。

さて、文献 [10] は整理した遷転情報から官職の遷転にはどのようなケースが存在し、そこにはいかなる傾向が見られたか、を分析している。これこそが同書の主題なのであるが、そこにもいくつかの問題が見受けられる。まず、唐代約三百年を初唐・中唐・晩唐の三期に区分した点が挙げられる。もとより、初唐・中唐・晩唐という時期区分は、歴史学上の意味を有するものである。ただ、対象とする官職全てにおいて、この三期ごとに遷転の傾向を考えようとするのは、やや乱暴にも思われる。

ある人物が官職 A に着任する前はどのような官職に就いていたか、またそこからいかなる官職に遷ったか、それが同書の遷移情報を分析する基本の姿勢である。従って、官職 A を中心としてその直前・直後の官職がどうであったか、ということしか考慮されていない憾みがある。遷転の傾向を考察するためには、直前に在職したポストだけでなく、前後にどのような官職を経験していたか、そのことについても考察を進める必要がある。

以下、本節で述べた問題点についてまとめておく。

### 遷転情報の整理に関する問題

1. 刊行後、三十年近くを経て、その間、営々と積み重ねられてきた各官職ごとの遷転情報整理の成果が生かされていない。
2. 使用した史料が根本的なものに限られる。
3. 対象となる官職が中央の重要な文官にのみ限られている。

4. 情報に対する解釈に、やや恣意的なところが見られる。

### 整理した情報の分析に関する問題

1. 唐代約三百年を初唐・中唐・晩唐の三期に区分して、対象とする官職全てにおいて、この三期ごとに遷転の傾向を考えようとする。
2. ある官職に注目したとき、その直前・直後にある人物がどの官職に就いていたか、ということしか考慮されていない。

以下の節では関連ルールによる分析によって、これらの問題点をどのように解決できるか、或いは未解決として残されるかについて述べていく。

## 4 関連ルールマイニング

本節では任官の傾向を示す関連ルールと、その関連ルールを抽出する方法である関連ルールマイニングについて、その概要を説明する<sup>11</sup>。

### 4.1 関連ルール

唐代の官僚制に存在した各々の官職をアイテム (item)、ある一人の人物が任官した官職の集合 (アイテムセット) をトランザクション (transaction) と呼ぶ。このトランザクションの全てを集めたものがデータベースである。このデータベースを解析すると、例えば、「中書舎人についたことのある人物はしばしば、中書侍郎についている」というような知識を得ることがある。このような情報を以下のような形で表現したものが関連ルール (association rule) である。官職  $P_0, P_1, \dots, P_n$  について

$$P_0 \Leftarrow P_1, \dots, P_n (M, N).$$

このルールは「官職  $P_1, \dots, P_n$  についたことのある人物はしばしば、官職  $P_0$  についている」という任官の傾向を示している。本稿では一つの官職への任官の傾向に注目したいため、関連ルールの結論部には官職を一つしか許さないように制限している (上記のルールでは  $P_0$  が結論)。  $M, N$  の意味については後述。

<sup>11</sup>但し、定義は一般のものではなく、本稿で必要とするものに限定してある。関連ルールマイニングについての参考文献は非常に多いが、文献 [18], [22] のみをあげるにとどめる。用語については文献 [22] に従う。

## 4.2 相関ルールのサポート、確信度

我々がデータベースを分析するとき、有用な相関ルールのみを抽出したい。この有用さを表現する指標として、相関ルールのサポート (support)、確信度 (confidence) がある。上記の相関ルールに即してこれを述べると、相関ルールのサポートは  $M\%$  であり<sup>12</sup>、これは官職  $P_1, \dots, P_n$  に任官した人全体に対する割合を示す。また、相関ルールの確信度は  $N\%$  であり、官職  $P_1, \dots, P_n$  に任官した人が官職  $P_0$  につく割合である。ここで官職の並びは任官の時間的な順序を意味しない。

相関ルールの最小サポート、最小確信度を与えて、これらを充たす相関ルールを求めるアルゴリズムとして、Apriori が提出されている [18]。

## 5 分析の方法と結果

### 5.1 分析の対象

文献 [10] による官人の遷転情報を対象とする。同書に様々な問題があることは、節 3 で述べたとおりである。ここであえて同書を用いる理由としては、我々が今後、独自に任官情報を収集・分析する過程で、どのような点に留意すべきか、あらかじめ見通しを得ておくということが挙げられる。対象とするデータは中国唐代の官人 1250 人<sup>13</sup>が任官した官職名<sup>14</sup>とその時期 (当時在位していた皇帝) である。10000 件ほどの任官情報である。また、各任官情報を時期によって分類し、その時期ごとにデータベースを作成した。時期の分類は文献 [10] により、初唐・中唐・晩唐とした。各時期は在位していた皇帝により、初唐は高祖から玄宗まで (618-756)、中唐は肅宗から敬宗まで (757-826)、晩唐は文宗から哀帝まで (827-907) とする。

### 5.2 分析の方法

アルゴリズム Apriori の実装として、本稿では Borgelt によるものを用いた ([21],[22])。

文献 [10] の官職遷転表のデータ化には YAML フォーマット [20] を採用した。データ化された官職遷転表からトランザクションを抽出し、PMML (Predictive

Model Markup Language) フォーマットに変換する<sup>15</sup>。Borgelt の Apriori の実装を使用して得られた相関ルールのマイニング結果も、PMML フォーマットで保存する<sup>16</sup>。

なお官職遷転表からトランザクションを作成する方法は以下である。まず一つの人名に注目し、一つのトランザクションを作成する。その人名を含む「官職 X 遷入表」あるいは「官職 X 遷出表」があり、官職 A のところにチェックがあるとする。このとき、このトランザクションに官職 X、官職 A を含める。この手順を全ての人名、全て官職遷転表について行なう。

### 5.3 抽出されたルール

#### 唐代全体: 確信度 50%以上のルール

員外郎 <= 殿中侍御史 (10.2, 63.0)  
員外郎 <= 郎中 (27.9, 62.2)  
員外郎 <= 侍御史 (10.9, 58.1)  
員外郎 <= 監察御史 (13.5, 58.0)  
郎中 <= 員外郎 (29.8, 58.3)  
郎中 <= 侍御史 (10.9, 55.9)  
郎中 <= 諫議大夫 (10.3, 51.2)  
郎中 <= 殿中侍御史 (10.2, 51.2)  
郎中 <= 中書舍人, 員外郎 (11.4, 60.6)  
中書舍人 <= 中書侍郎 (15.2, 51.1)

#### 初唐: 確信度 40%以上のルール

員外郎 <= 郎中 (11.9, 41.8)  
中書舍人 <= 吏部侍郎 (12.1, 51.5)  
中書舍人 <= 中書侍郎 (13.0, 43.8)  
中書侍郎 <= 中書令 (10.5, 49.2)

#### 中唐: 確信度 70%以上のルール

員外郎 <= 補闕 (11.3, 86.5)  
員外郎 <= 監察御史 (17.4, 76.2)  
員外郎 <= 侍御史 (14.2, 72.3)  
員外郎 <= 殿中侍御史 (14.6, 71.6)  
員外郎 <= 拾遺 (12.0, 70.9)  
員外郎 <= 中書舍人, 郎中 (10.5, 83.3)  
郎中 <= 中書舍人, 員外郎 (10.9, 80.0)

<sup>12</sup>ここでのサポートの定義は文献 [18] ではなく、文献 [22] による。

<sup>13</sup>資料の性格から多くはかなりの高位に達した者。

<sup>14</sup>文献 [10] の研究対象である中央政府の、それも中枢部に比較的近い文官に概ね限定される。

<sup>15</sup>PMML はマイニングの対象データと結果を表現するデータマイニング記述言語で、その仕様が XML のドキュメントとして Data Mining Group により策定されている。文献 [23] を参照。

<sup>16</sup>将来的にこのフォーマットをデータ変換に使用する予定。

## 晩唐: 確信度 70%以上のルール

員外郎 <= 中書侍郎, 中書舍人 (10.8, 70.0)  
中書舍人 <= 禮部侍郎 (12.5, 78.3)  
中書舍人 <= 戸部侍郎, 員外郎 (13.0, 75.0)  
中書舍人 <= 中書侍郎, 兵部侍郎 (10.0, 70.3)  
中書侍郎 <= 門下侍郎 (12.2, 84.4)  
戸部侍郎 <= 中書侍郎, 中書舍人 (10.8, 75.0)  
戸部侍郎 <= 門下侍郎, 中書侍郎 (10.3, 73.7)

## 5.4 分析の結果とその解釈

以上の抽出されたルールから以下のような中国学的な解釈ができる。

- 「八儁」と呼ばれるエリートコース (文献 [11]) が、実際に時代をおって確立していく様子が見取れる。すなわち、「八儁」に含まれる官職を歴任したケースの割合が 6、7 割と極めて高くなる。
- 時代が下る (初唐から晩唐へ) に従って、確信度の高いルールが抽出できる。しかし、晩唐全体でどれだけのデータがあるのか不明のため、記録の不足は否めない。黄巢の乱 (唐の政府に対する大規模な反乱、875-884) をはじめ唐末・五代 (五代は唐の滅亡から宋が建国するまでの混乱期、907-960) の戦禍で晩唐の公式記録は編纂されなかった、或いはされても不充分、もしくは散逸した例が多いことがこのような傾向を示している可能性がある。
- 初唐は確信度の大きいルールが抽出しづらい。ルールが確立する段階が、即ち高祖、太宗期 (618-649)、より範囲を広げれば武則天期 (中宗・睿宗の時期を含む、684-705) だった可能性がある。つまり本当にルールがまだ後の時代ほどには出来上がっていなかったことを示す。
- 中唐以後は員外郎が遷転に重要な位置を占めるようになる。特に拾遺、補闕、侍御史、殿中侍御史、監察御史との関りが強い。これらの官職がエリートであることはいうまでもないが、次のような記録も見られる (9 世紀前半の随筆『唐国史補』巻下)。

(複数いる) 宰相は (互いを) 「元老」といったり、「堂老」といったりする。  
(中書・門下) 両省は「閣老」と呼び、尚書省の左右丞・侍郎・郎中は「曹長」と

呼ぶ。員外郎・御史・拾遺・補闕は「院長」と呼ぶ。上のもので下のものを兼ねさせることはできて、下のもので上のものを兼ねさせることはできないのだが、ただ侍御史のみは「端公」と呼ぶ。<sup>17</sup>

品階の差を超えて、互いに、また他の職にある者でも、「員外郎・御史 (殿中侍御史・監察御史)・拾遺・補闕」には同じ「院長」の敬称を用いていたという事実も、員外郎とその他の官職の関わりが強かったことを示す可能性がある。

- 中書舍人については以下のようなルールを抽出できる。ここから中唐では他の時期に比べて、員外郎、郎中との結び付きが強いことが分る。(中唐では中書舍人の経験者は 6、7 割が員外郎、郎中経験者である。他の時期は 4 割程度以下。但し、初唐はある程度の確信度で以下の形のルールは抽出できない。)

唐全体: 郎中 <= 中書舍人 (23.9, 39.8)  
員外郎 <= 中書舍人 (23.9, 47.5)

中唐: 郎中 <= 中書舍人 (15.7, 66.7)  
員外郎 <= 中書舍人 (15.7, 69.4)

晩唐: 郎中 <= 中書舍人 (35.0, 40.3)  
員外郎 <= 中書舍人 (35.0, 43.4)

## 6 結論及び課題

### 6.1 結論

節 3 で述べた文献 [10] の問題点についての結論を述べる。整理した情報の分析に関する問題について、本稿では遷転時期の順序によらないルールを抽出しているため、時期の離れた官職の間の遷転の傾向がみてとれる。但し、遷転情報の整理に関する問題については、同書の官職遷転表をデータとして使用したため、これらの問題については未解決である。また時代区分の恣意性についても、同書の時代区分を採用したためならの発展を見ない。これらの課題については以下で取り上げる。

<sup>17</sup> 文献 [14] 頁 49 参照。原文「宰相相呼為元老、或曰堂老。兩省相呼為閣老、尚書丞郎郎中相呼為曹長。外郎御史遺補相呼為院長。上可兼下、下不可兼上、唯侍御史相呼為端公。」

## 6.2 今後の課題

中国唐代の官僚が遷転した様子を解明しようとする時に、本稿が残した今後の課題として以下が考えられる。

- 時系列パターンの抽出<sup>18</sup>
- 時代区分の問題
- 科挙、武官、地方官、流外官を対象とすること

データベースから任官の時間的な順序を考慮した時系列パターンを抽出するためには、時期情報の順序を扱えるように歴史的な時間というものをモデル化する必要がある。時系列パターンを抽出することによって、より詳しい官僚の遷転の様子を解明することが期待できる。たとえば、病坊<sup>19</sup>の解明などがあげられる。

本稿の分析では文献 [10] において行われていた初唐・中唐・晩唐の三区分を利用している。分析に使用できるデータ数が少ないため、時代傾向を知る上で三区分程度の時代区分をするしかないことが理由である。このような恣意的な三区分ではなく、抽出される相関ルールが依存するような時代区分を発見することができるかどうかという問題がある。これには現在の分析対象となっている任官情報よりも多くの情報を必要とする。

本稿の分析対象のデータは文献 [10] の官職遷転表にもとづく。これは書名が示すように中央重要文官のみを対象としていて、科挙、武官、地方官、流外官を扱っていない。相関ルールマイニングの対象をこれらの官職に拡大することによって、「エリートコース」以外の官職遷転のパターンを抽出することができると期待できる。

## 7 おわりに

現在、我々は漢字文献の知識ベース化を試みるプロジェクトを推進中である [1]。そのサブプロジェクト唐代官職知識ベースにおいて、官職に関係する各種文献の任官情報をデータ化している。現在、人名 14000、任官情報 54000 件のデータ化が完了している。これらの任官情報は文献 [10] のデータ化と同様に YAML フォー

<sup>18</sup>列のパターンを扱うアルゴリズムについては文献 [19] を参照。

<sup>19</sup>例えば、『両京新記』には、宰相の激務にたえられなかった者は書物の筆写・収蔵・校訂を主管するだけの部署である秘書省の長官・秘書監に選ばれることがあった。このため、秘書監は宰相の「病坊」（療養所）と呼ばれた、とある。文献 [15] 頁 1404、1405 を参照。

マット [20] で書かれている。但し、データモデルについては各文献ごとに異なる<sup>20</sup>。

データ化の終了した大量データを分析するにあたっては、以下の問題を解決する必要がある。データ化の終了している任官の情報に対して、人物を同定するための作業を行わなければならない。すなわち、同姓同名の人物の弁別、別名による人物の同定等である。また、文献ごとに異なる官職名を同定したり、官職をクラス分けして処理するためのクラス階層化の作業も必要である。

これらのデータを分析の対象とすることによって、節 3 で述べた文献 [10] 『唐代中央重要文官遷転途徑研究』の問題点のうち、本稿で積み残した問題を解決することが期待できる。また、時代区分の問題、及び科挙・武官・地方官・流外官を対象とすることも可能である。

本稿の応用として、時期の分類や官職のクラス分けなどのパラメータを指定してマイニングするために、相関ルールのマイニングアルゴリズムを Web システムに組み込むことが考えられる。また、本稿は官職遷転の傾向を調べるために、相関ルールマイニングを用いた訳であるが、その他の方法論、例えば学習、帰納論理プログラミングなどの方法論との比較が考えられる。

**謝辞** 本稿を書くにあたり、以下の方々にお世話になりました。深く感謝いたします。安岡孝一さんには唐代知識ベースのプロジェクトにおいてお世話になっております。秋山陽一郎さん、牛根靖裕さん、山田崇仁さん、森華さんには原稿に詳しいコメントを戴きました。官職知識ベースのプロジェクトメンバーである橋本英治さん、森華さん、大井留美さん、鈴鹿摩耶さんにはデータの作成にご協力いただきました。また、第一筆者は特に以下の方々にお世話になりました。橋本英治さん、森華さんには『唐代中央重要文官遷転途徑研究』について多くのことをご教示いただきました。秋山陽一郎さんには日頃より人文学、計算機科学についての議論を通して多くの影響を受けております。最後に概要論文に対してコメントいただいた匿名査読者の皆様に感謝いたします。

## 参考文献

- [1] 秋山陽一郎, 白須裕之, 永田知之: 中国古典学知識ベースにおける信頼性評価モデルの一試案, 第 17 回 東洋学へのコンピューター利用, 2006.

<sup>20</sup>任官情報の概念モデルについては文献 [7] を参照。

- [2] 郁賢皓, 胡可先: 唐九卿考, 中国社会科学出版社, 2003.
- [3] 池田温: 韓琬『御史台記』について, 布目潮風博士古稀記念論集東アジアの法と社会, pp.111-138, 汲古書院, 1990.
- [4] 池田温: 律令官制の形成, 岩波講座世界歴史 5, 岩波書店, pp.277-323, 1970.
- [5] 歐陽修, 宋祁: 新唐書, 中華書局, 1975.
- [6] 嚴耕望: 唐僕尚丞郎表, 中央研究院歷史語言研究所, 1956.
- [7] 白須裕之: 唐代任官情報の概念モデル — 時間、文獻に依存する情報のための分析パターン —, 情報処理学会研究報告, 2006-CH-72, 2006.
- [8] 謝枋得輯, 吳道南補輯: 新鐫饗饗必用增補秘笈新書, 莊嚴文化事業有限公司, 1997.
- [9] 曾慥: 類說, 文学古籍刊行社, 1955.
- [10] 孫国棟: 唐代中央重要文官遷轉途徑研究, 龍門書店, 1978.
- [11] 礪波護: 唐代の県尉, 唐の行政機構と官僚, 中公文庫, pp.73-122, 1998.
- [12] 福山敏男: 校注 兩京新記卷第三及び解説, 福山敏男著作集 6, 中央公論美術出版, pp.105-184, 1983.
- [13] 封演撰, 趙貞信校注: 封氏聞見記校注, 中華書局, 1958.
- [14] 李肇: 唐国史補, 古典文学出版社, 1957.
- [15] 李昉等: 太平広記, 中華書局, 1961.
- [16] 李林甫等撰, 陳仲夫点校: 唐六典, 中華書局, 1992.
- [17] 劉昫等: 旧唐書, 中華書局, 1975.
- [18] R. Agrawal, T. Imielinski and A. Swami, Mining association rules between sets of items in large databases, Proc. SIGMOD, 207-216, 1993.
- [19] R. Agrawal and R. Srikant, Mining Sequential Patterns, Proc. ICDE, 3-14, IEEE, 1995.
- [20] O. Ben-kiki, C. Evans and B. Ingerson, YAML Ain't Markup Language(YAML™) Version 1.1, 2004.
- [21] C. Borgelt, Efficient Implementations of Apriori and Eclat, Workshop of Frequent Item Set Mining Implementations, 2003.
- [22] C. Borgelt, Apriori - Association Rule Induction/Frequent Item Set Mining, <http://fuzzy.cs.uni-magdeburg.de/~borgelt>, 2005.
- [23] Data Mining Group, PMML 3.1 Specification, <http://www.dmg.org/pmml-v3-1.html>, 2006.