

А.А.Давыдов

(Москва)

СТОХАСТИЧЕСКАЯ ДИНАМИКА РАЗВИТИЯ ЧЕЛОВЕКА В СОЦИУМЕ: МОДЕЛЬ ДВОЙНОЙ ЦЕПИ МАРКОВА

В статье представлена модель двойной цепи Маркова первого порядка с двумя скрытыми состояниями, которая удовлетворительно описывает стохастическую динамику развития человека в 90 странах мира, сгруппированных в три группы (Высокий уровень развития человека, Средний уровень развития человека, Низкий уровень развития человека), измеренного с помощью Human Development Index (HDI) за период 1980-2010 гг.

The article presents the Double Chain Markov Model (DCMM), Number of hidden states = 2, Order of the hidden model = 1, Order of the visible model = 1, which satisfactorily describes the stochastic dynamics of Human Development in 90 countries, grouped into three groups (High Human Development, Medium Human Development, Low Human Development), measured using the Human Development Index (HDI) for the period 1980-2010 years.

Ключевые слова: системная социология, социум, стохастическая динамика, развитие человека, двойная цепь Маркова.

Введение

В международной деятельности Организации Объединенных Наций (ООН) [1] фундаментальной социальной концепцией и приоритетной целью международного сообщества является «Human Development» (развитие человека) в социуме (множество стран мира), которая базируется на Всеобщей декларации прав человека, принятой Генеральной Ассамблеей ООН в 1948 году. В сравнительных международных исследованиях ООН, развитие человека в странах мира измеряют с помощью индекса Human Development Index (HDI) – Индекса развития человеческого потенциала (ИРЧП)[2].

Андрей Александрович Давыдов – доктор философских наук, главный научный сотрудник Института социологии РАН. E-mail: keyly@yandex.ru

Для удобства аналитической и практической деятельности, ООН использует следующую классификацию стран мира по значению HDI: Very High Human Development (очень высокий уровень развития человека), High Human Development (высокий уровень развития человека), Medium Human Development (средний уровень развития человека), Low Human Development (низкий уровень развития человека). По данной классификации ООН [2], Россия в 2010 г. входила в группу стран мира High Human Development (высокий уровень развития человека).

В 2010 г. для классификации стран мира по уровню Human Development Index (HDI) использовались квартили [2,с.139], а в предыдущие годы классификация была основана на предварительно установленных точках разделения значений HDI, что затрудняет сопоставимость результатов в различные моменты времени.

Известно [2], что в зависимости от множества внутренних факторов (демографических, экономических, политических, военных и т.д.) с течением времени происходит переход стран мира из одной группы Human Development в другую группу или страна мира длительное время находится в одной из групп Human Development. Вместе с тем, недостаточно изученным является вопрос о стохастических закономерностях переходов стран мира между группами на уровне социума в целом. В этой связи автор поставил следующую исследовательскую задачу. Выявить стохастические закономерности динамики перехода стран мира между группами Human Development на уровне социума в целом.

Поставленная исследовательская задача является одной из классических задач системной социологии [3], решение которой актуально для развития теории системной динамики социума, последующего проведения компьютерного моделирования и прогнозирования динамики социума, практических приложений в деятельности ООН.

Теория

Решение поставленной исследовательской задачи осуществлялось в рамках системной социологии [3] на основе классической теории динамики стохастических систем, которая давно и широко [4-6] используется для анализа динамики социальных систем. Данное исследование базировалось на теоретическом предположении, которое имеет солидное эмпирическое подкрепление [3], согласно которому на уровне социума в целом (множество стран мира) действуют общесистемные стохастические закономерности динамики.

Методология

В исследовании использовалась естественнонаучная методологическая парадигма системной социологии [3], в частности, методология анализа цепей Маркова, которая является одной из классических методологий [4-6] анализа и моделирования стохастической динамики социальных систем в системной социологии.

Методика

Были использованы значения Human Development Index (HDI) для 169 стран мира по пятилетним интервалам за период 1980-2010 гг. из базы данных ООН [1]. Поскольку методика выделения классов, используемая ООН [2,с.139], не позволяет сравнивать группы за период 1980-2010 гг., и вызывает обоснованные методические сомнения, то автором была использована следующая формальная вычислительная процедура. Для выявления трех групп стран мира по уровню значений Human Development Index (HDI) использовался кластерный анализ по методу K-Means из пакета SPSS с принудительным выделением трех групп, а именно, Высокий уровень HDI, Средний уровень HDI и Низкий уровень HDI. Выделение трех классов было обусловлено незначительным числом стран мира для анализа переходных вероятностей, а также для обеспечения сопоставимости с другими известными классификациями стран мира, например, классификацией И.Валлерстайна [7], который выделял «Ядро», «Полупериферию» и «Периферию» в Мир-системе (социуме). Переходные вероятности между выделенными классами по пятилетним интервалам времени (стохастические матрицы) вычислялись с помощью пакета SPSS.

Для статистического анализа динамики распределения стран мира за период 1980-2010 гг. по группам и стохастических матриц переходных вероятностей между группами стран мира, было выделено 90 стран мира, по которым имелись значения классов без пропусков наблюдений за период 1980-2010 гг.

Анализ осуществлялся с помощью пакета March 3.0 [8], разработанный A.Berchtold. В пакете March 3.0 [8] реализованы следующие модели классических, обобщенных и комбинированных цепей Маркова: Independence model, First-order homogeneous Markov chain, High-order homogeneous Markov chain, Mixture Transition Distribution (MTD) model, Hidden Markov Model (HMM), Double Chain Markov Model (DCMM). Вычисление значений параметров моделей осуществляется с помощью комбинации алгоритмов Expectation-Maximization algorithm (EM), Genetic algorithm (GA), Best Improving Step (BIS). Селекция моделей осуществляется по критериям Log-likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC). Заинтересованный читатель может

более детально ознакомиться с математическим описанием моделей, алгоритмами оценивания значений параметров моделей и селекции моделей в справочной системе пакета March [8].

Полученные результаты

В Приложении 1 представлена динамика стран мира, сгруппированных по уровню значений Human Development Index (HDI). В Приложении 2 представлены переходные вероятности между выявленными группами стран мира по пятилетним интервалам времени, которые свидетельствуют, что переходные вероятности зависят от времени. Это означает, что цепь Маркова неоднородна по времени.

Генерирование множество моделей High-order homogeneous Markov chain, Mixture Transition Distribution (MTD) model, Hidden Markov Model (HMM), Double Chain Markov Model (DCMM), реализованных в пакете March 3.0 [8] и их последующая селекция по вычислительным критериям аппроксимации данных, а также содержательным и иным критериям системной социологии [3], в частности, простоты модели, показали, что наилучшей является Double Chain Markov Model (DCMM).

Double Chain Markov Model (DCMM) является неоднородной цепью Маркова (переходные вероятности зависят от времени) и обобщает модели Hidden Markov Model (HMM) и High-order Markov Chain Model. В качестве наглядной иллюстрации, на рис. 1 представлена Double Chain Markov Model (DCMM) первого порядка. Порядок модели (Order of the model) соответствует лагу (запаздыванию). Например, Order of the visible model = 1 соответствует классической цепи Маркова, в которой Y_t зависит только от Y_{t-1} и не зависит от более ранних моментов времени $Y_{t-2}, Y_{t-3}, \dots, Y_{t-n}$.

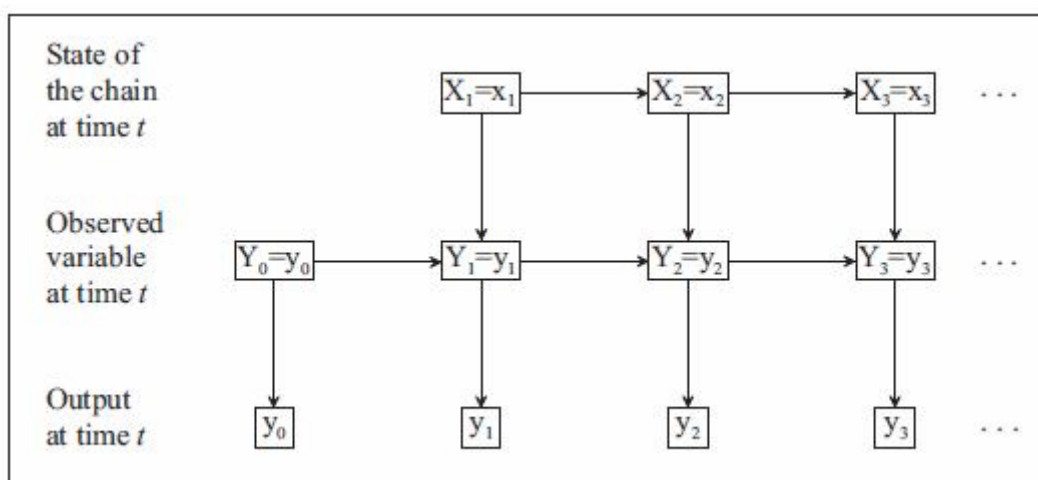


Рис.1. Double Chain Markov Model (DCMM)

Double Chain Markov Model (DCMM) описывается следующими параметрами:

- Множество скрытых состояний (hidden states). $S(X) = \{1, \dots, M\}$.
- Множество наблюдаемых выходов (possible outputs). $S(Y) = \{1, \dots, K\}$.
- Распределение вероятностей первого скрытого состояния. $\pi = (\pi_1, \dots, \pi_M)$.
- Матрица переходных вероятностей между скрытыми состояниями. $A = \{a_{ij}\}$, $i, j \in S(X)$.
- Множество матриц вероятностей переходов последовательных выходов Y_t данного конкретного состояния X_t . $C = \{c_{ijs}\}$, $i, j \in S(Y)$, $s \in S(X)$.
 $C = \{C^{(s)}\}$ с $C^{(s)} = [c_{ij}^{(s)}]$
- $\mu = \{\pi, A, C\}$

Double Chain Markov Model (DCMM) разработана A.Berchtold [8] для анализа и моделирования неоднородных цепей Маркова, где глобальные скрытые состояния определяют различия между стохастическими матрицами переходных вероятностей наблюдаемых состояний системы. В содержательном смысле, Double Chain Markov Model (DCMM) адекватна известным закономерностям динамики социума [2-3], в частности, известно, что на динамику значений Human Development Index (HDI) влияют, как локальные внутренние национальные факторы (экономические политические, военные и т.д.) специфичные для конкретной страны мира в каждый момент времени, так и глобальные закономерности динамики социума, например, длинные волны И.Шумпетера [9], в результате которых социум переходит из одного состояния в другое, например, стабильное состояние – нестабильное состояние, в частности, глобальные экономические, энергетические, экологические и т.д. кризисы, мировые войны. Однако, в ряде случаев, в частности, на периоде 1980-2010 гг., трудно измерить и содержательно проинтерпретировать скрытые состояния социума и переходные вероятности между данными скрытыми состояниями, а Double Chain Markov Model (DCMM) позволяет решить данную сложную задачу. В целом, Double Chain Markov Model (DCMM) является адекватной и плодотворной моделью для анализа стохастической динамики социума.

В результате проведенной селекции моделей была выбрана Double Chain Markov Model (DCMM), Number of hidden states = 2, Order of the hidden model = 1, Order of the visible model = 1. Ниже представлены полученные результаты, выдаваемые в пакете March 3.0 [8].

```
*****
* Double Chain Markov Model *
*****
```

Number of independent sequences used for the computation: 90

Number of possible outputs of the dependent variable: 3

Number of hidden states: 2

Order of the hidden model: 1

Order of the visible model: 1

Number of independent parameters: 11

Number of components in the log-likelihood: 540

Overall precision of the estimation (type I error = 0.05): 0.155929

Log-likelihood: -119.618

AIC: 261.236

BIC: 308.443

Type I error for confidence intervals: 0.05

*** Hidden model ***

First hidden state distribution (last column = half confidence interval):

1.0000	0.0000		0.1190
--------	--------	--	--------

Combinations of values corresponding to each row of the transition matrix:
row number, first lag, second lag, ..., last lag of the hidden variable

1		1
2		2

Transition matrix of the hidden model (reduced form, last column = half confidence interval):

1		0.7313	0.2687		0.0687
2		0.0000	1.0000		0.0841

* Analysis of the transition matrix of the Markov chain *

The chain is positive regular with only one (final) non-cyclical class.

There are no transient states.

There are no absorbing states.

States belonging to final class 1 (non-cyclical):

1	2
---	---

The limit distribution of the Markov chain is

0.0000	1.0000
--------	--------

*** Visible model ***

* Transitions between successive observations, state 1 *

Combinations of values corresponding to each row of the transition matrix:
row number, first lag, second lag, ..., last lag of the observed variable

1		1
2		2
3		3

Transition matrix of the visible model (last column = half confidence interval):

1		0.9511	0.0489	0.0000		0.1349
2		0.0000	0.9515	0.0485		0.1303
3		0.0000	0.0662	0.9338		0.1559

* Analysis of the transition matrix of the Markov chain *

The chain is positive regular with only one (final) non-cyclical class.

There are no transient states.

There are no absorbing states.

States belonging to final class 1 (non-cyclical):

1	2	3
---	---	---

The limit distribution of the Markov chain is

0.0000	0.5768	0.4232
--------	--------	--------

* Transitions between successive observations, state 2 *

Transition matrix of the visible model (last column = half confidence interval):

1		0.9928	0.0072	0.0000		0.1134
2		0.1628	0.8372	0.0000		0.1217
3		0.0000	0.0288	0.9712		0.1372

* Analysis of the transition matrix of the Markov chain *

The chain is positive regular with only one (final) non-cyclical class.

There are no transient states.

There are no absorbing states.

States belonging to final class 1 (non-cyclical):

1	2	3
---	---	---

The limit distribution of the Markov chain is

0.9575	0.0425	0.0000
--------	--------	--------

*** Hidden states ***

Analysis:

Crosstable between observations (rows) and hidden states (columns):

39	176
136	47
40	102

Chi-square statistic: 141.1

The independence hypothesis is rejected at the 95% level (df=2).

Cramer's V statistic: 0.511171

Theil's U statistic with observations dependent: 0.123837

Theil's U statistic with hidden states dependent: 0.1998

Обсуждение полученных результатов

Стохастические матрицы переходных вероятностей, представленные в Приложении 2, свидетельствуют, что значения переходных вероятностей зависят от времени и диагональные элементы имеют наибольшее значение. Это означает, что стохастический процесс перехода между группами стран мира на уровне социума является неоднородным (зависит от времени), а нахождение в группе наиболее устойчиво с течением времени. Полученные результаты хорошо согласуются с другими результатами [10] использования цепей Маркова для других переменных для групп стран мира на уровне социума в целом. Следовательно, имеются эмпирические основания предполагать, что наблюдается общая стохастическая закономерность на уровне социума в целом.

Исходя из значения переходных вероятностей в скрытых состояниях (см. выше Transitions between successive observations, state 1 и Transitions between successive observations, state 2), а также предельных распределений в скрытых состояниях (см. выше The limit distribution of the Markov chain is), скрытые состояния можно содержательно интерпретировать следующим образом. Первое скрытое состояние социума - неблагоприятно для Human Development (развития человека), а второе скрытое состояние социума благоприятно для Human Development (развития человека).

Выводы

В результате проведенного исследования имеются эмпирические основания полагать, что переход 90 стран мира между тремя группами по уровню Human Development (развития человека), измеренного с помощью Human Development Index (HDI) за период 1980-2010 гг., может быть удовлетворительно описан с помощью модели двойной цепи Маркова первого порядка с двумя скрытыми состояниями социума.

Благодарность. Автор признателен André Berchtold - Ph.D., Professor, Université de Lausanne, Institut de Mathématiques Appliquées, Switzerland, автору модели Double Chain Markov Model (DCMM), разработчику пакета March 3.0, за помощь в анализе данных.

ЛИТЕРАТУРА

1. United Nations Human Development Programme. (<http://hdr.undp.org/en/>)
2. Human Development Report 2010. The Real Wealth of Nations: Pathways to Human Development. UN, 2010. Доклад о развитии человека 2010. Реальное богатство народов: пути к развитию человека. М.: «Весь мир», 2010. (<http://hdr.undp.org/en/>)
3. Давыдов А.А. Системная социология: введение в анализ динамики социума. М.: ЛКИ, 2007.
4. Бартоломью Д. Стохастические модели социальных процессов: Пер. с англ. М.: Финансы и статистика, 1985.
5. Робертс Ф.С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам: Пер. с англ. - М.: Наука. Гл. ред. физ.-мат. лит., 1986.
6. Helbing D. Quantitative Sociodynamics: Stochastic Methods and Models of Social Interaction Processes. N.Y.: Springer, 2010.
7. Wallerstein I. The Modern World System. V.I-III. N.Y.: Plenum Press, 1974-1989.
8. March 3.0. (<http://www.andreberchtold.com/march.html>)
9. Давыдов А.А. Волны инноваций и числа Фибоначчи: оценка перспективности гипотезы. Официальный сайт РОС, 2010. (http://www.ssa-rss.ru/index.php?page_id=22&id=53)
10. Давыдов А.А. Модернизация России, полезный опыт Китая и теория сложных систем. Официальный сайт РОС, 2010. (http://www.ssa-rss.ru/index.php?page_id=22&id=53)

ПРИЛОЖЕНИЯ

Приложение 1. Распределение стран мира по уровню HDI,% (кластерный анализ K-Means).

Уровень HDI/год	1980	1985	1990	1995	2000	2005	2010
Высокий уровень HDI	41,0	37,9	36,5	31,0	32,1	29,6	47,3
Средний уровень HDI	31,6	37,9	38,1	43,4	39,4	40,2	27,8
Низкий уровень HDI	27,4	24,2	25,4	25,6	28,5	30,2	24,9
Количество стран мира	95	103	118	129	137	169	169

Приложение 2. Стохастические матрицы переходных вероятностей между группами стран мира,% по пятилетним интервалам времени за период 1980-2010 гг.

Cluster Number of Case_1980 * Cluster Number of Case_1985 Crosstabulation

% within Cluster Number of Case_1980

		Cluster Number of Case_1985		
		Высокий уровень HDI	Средний уровень HDI	Низкий уровень HDI
Cluster Number of Case_1980	Высокий уровень HDI	94,9%	5,1%	
	Средний уровень HDI		100,0%	
	Низкий уровень HDI		12,0%	88,0%

Cluster Number of Case_1985 * Cluster Number of Case_1990 Crosstabulation

% within Cluster Number of Case_1985

		Cluster Number of Case_1990		
		Высокий уровень HDI	Средний уровень HDI	Низкий уровень HDI
Cluster Number of Case_1985	Высокий уровень HDI	94,7%	5,3%	
	Средний уровень HDI		92,3%	7,7%
	Низкий уровень HDI			100,0%

Cluster Number of Case_1990 * Cluster Number of Case_1995 Crosstabulation

% within Cluster Number of Case_1990

		Cluster Number of Case_1995		
		Высокий уровень HDI	Средний уровень HDI	Низкий уровень HDI
Cluster Number of Case_1990	Высокий уровень HDI	81,4%	18,6%	
	Средний уровень HDI		97,8%	2,2%
	Низкий уровень HDI			100,0%

Cluster Number of Case_1995 * Cluster Number of Case_2000 Crosstabulation

% within Cluster Number of Case_1995

		Cluster Number of Case_2000		
		Высокий уровень HDI	Средний уровень HDI	Низкий уровень HDI
Cluster Number of Case_1995	Высокий уровень HDI	100,0%		
	Средний уровень HDI	7,5%	90,6%	1,9%
	Низкий уровень HDI		6,3%	93,8%

Cluster Number of Case_2000 * Cluster Number of Case_2005 Crosstabulation

% within Cluster Number of Case_2000

		Cluster Number of Case_2005		
		Высокий уровень HDI	Средний уровень HDI	Низкий уровень HDI
Cluster Number of Case_2000	Высокий уровень HDI	97,7%	2,3%	
	Средний уровень HDI		98,1%	1,9%
	Низкий уровень HDI		2,6%	97,4%

Cluster Number of Case_2005 * Cluster Number of Case_2010 Crosstabulation

% within Cluster Number of Case_2005

		Cluster Number of Case_2010		
		Высокий уровень HDI	Средний уровень HDI	Низкий уровень HDI
Cluster Number of Case_2005	Высокий уровень HDI	100,0%		
	Средний уровень HDI	44,1%	55,9%	
	Низкий уровень HDI		17,6%	82,4%