

01

Реконструкция архитектуры связей в цепочке из трех однонаправленно связанных систем методом причинности по Грейнджеру

© М.В. Корнилов^{1,2}, И.В. Сысоев^{1,2}

¹ Саратовский филиал Института радиотехники и электроники
им. В.А. Котельникова РАН, Саратов, Россия

² Саратовский национальный исследовательский государственный
университет им. Н.Г. Чернышевского, Саратов, Россия
E-mail: kornilovmv@gmail.com

Поступило в Редакцию 9 января 2018 г.

Предложена методика реконструкции связей в цепочке однонаправленно связанных осцилляторов на основе проведения трех тестов, оценивающих прямое и опосредованное взаимодействие методом условной причинности по Грейнджеру, с использованием прогностических моделей с полиномиальной нелинейностью. Показано, что данный подход позволяет более чем в 80% случаев верно определять архитектуру связей.

DOI: 10.21883/PJTF.2018.10.46103.17201

В настоящее время существуют различные подходы к детектированию направленной связанности между элементами сложных систем по их экспериментальным сигналам. Один из самых популярных — метод причинности по Грейнджеру [1]. Если подсистем больше двух, дополнительно к проблеме детектирования наличия и направления связей возникает проблема различения прямого и опосредованного взаимодействия. В частности, она представляет большой практический интерес для задач радиотехники, нейрофизиологии и климатологии. Общая идея решения данной проблемы известна: следует использовать прогностические модели, при построении которых учитывается информация не только от двух подсистем (ведомой и ведущей), но от нескольких (такой подход в [2] получил название условной причинности по Грейнджеру). Однако ее практическое воплощение, особенно для

нелинейных систем, оказывается затруднительным по многим причинам, главная из которых — недостаток объема имеющихся экспериментальных данных для надежной оценки всех коэффициентов [3]. Недостаточно высокая частота выборки, малая общая длина записи и, как следствие, малое доступное число колебаний и недоступность для измерения части подсистем не дают возможности решить проблему в общем, однако отдельные успешные частные решения возможны [4].

В настоящей работе рассматривается задача о реконструкции архитектуры связей в цепочке из трех однонаправленно связанных подсистем, демонстрирующих колебания с одним или несколькими основными временными масштабами, методом нелинейной причинности по Грейнджеру с использованием моделей в виде отображений последования с полиномиальными нелинейностями в соответствии с [2]. Цель работы — определить, насколько метод способен верно распознать архитектуру связей и, что наиболее существенно, различить прямое и опосредованное взаимодействие при разных уровнях связи (от очень слабых и не приводящих к существенным изменениям в динамике ведомых систем до достаточно сильных, ведущих к синхронизации).

Пусть в нашем распоряжении имеются скалярные временные ряды $\{x_n\}_{n=1}^N$, $\{y_n\}_{n=1}^N$ и $\{z_n\}_{n=1}^N$ от трех систем X , Y , Z длиной N отсчетов каждый. При этом система X считается ведомой и предполагается, что она не может воздействовать ни на Y , ни на Z . Система Y (промежуточная система) может воздействовать только на X , а система Z может воздействовать на X как непосредственно, так и опосредованно через Y или не воздействовать вообще (используя термин „цепочка“, мы в первую очередь подразумеваем, что не может быть обратных связей вида $X \rightarrow Y$ или $Y \rightarrow Z$ или $X \rightarrow Z$, т.е. если воздействие и распространяется, то в одну сторону). Всего, таким образом, возможно восемь случаев взаимодействия:

- 1) $Z \rightarrow Y \rightarrow X$ без непосредственного воздействия Z на X (опосредованная связь);
- 2) $Z \rightarrow X \wedge Z \rightarrow Y \rightarrow X$ — случай прямой и опосредованной связи одновременно;
- 3) $Z \rightarrow Y \wedge Z \rightarrow X$, но без воздействия Y на X , т.е. только прямая связь;
- 4) $Z \rightarrow X \wedge Y \rightarrow X$ — при отсутствии опосредованного взаимодействия;
- 5) отсутствие связи между Z и X , но наличие связи $Y \rightarrow X$;

6) $Z \rightarrow X$ — прямая связь, без искажения сигнала системы Y сигналом системы Z ;

7) отсутствие связи между Z и X , но наличие связи $Z \rightarrow Y$ (система X изолирована);

8) отсутствие каких-либо связей, все системы изолированы.

Такая постановка задачи может соответствовать системе беспроводных сенсоров, передающих информацию о напряжении свода здания или уровне задымления по цепочке [5], или соображениям анатомии, когда существуют восходящие связи от органов чувств к глубоким структурам мозга и далее в кору.

Поскольку тест на причинность по Грейнджеру может ответить только на вопрос о том, есть ли значимая связь между системами в заранее заданном направлении или нет, чтобы определить, какой из восьми рассмотренных случаев имеет место, нужно провести как минимум $\log_2 8 = 3$ тестирования на связанность. Для исследования связи $Z \rightarrow Y$, $Z \rightarrow X$ и $Y \rightarrow X$ наиболее очевидный подход состоит в построении прогностических моделей вида

$$\begin{aligned} x'_{n+\tau} &= f_x(\mathbf{x}_n), \quad x'_{n+\tau|Y} = f_{xy}(\mathbf{x}_n, \mathbf{y}_n), \quad x'_{n+\tau|Z} = f_{xz}(\mathbf{x}_n, \mathbf{z}_n), \\ y'_{n+\tau} &= f_y(\mathbf{y}_n), \quad y'_{n+\tau|Z} = f_{yz}(\mathbf{y}_n, \mathbf{z}_n) \end{aligned} \quad (1)$$

и дальнейшем расчете улучшения прогноза по формулам

$$PI_{y \rightarrow x} = 1 - \frac{\varepsilon_{xy}^2}{\varepsilon_x^2}, \quad PI_{z \rightarrow x} = 1 - \frac{\varepsilon_{xz}^2}{\varepsilon_x^2}, \quad PI_{z \rightarrow y} = 1 - \frac{\varepsilon_{yz}^2}{\varepsilon_y^2}. \quad (2)$$

Здесь \mathbf{x}_n , \mathbf{y}_n и \mathbf{z}_n — реконструированные методом временных задержек [6] по скалярным временным рядам векторы состояния, функции f_x , f_{xy} , f_{xz} , f_y и f_{yz} — полиномы общего вида, как предложено в [2], τ — дальность прогноза, x'_n и др. — предсказанные значения, соответствующие измеренным значениям x_n , PI — показатели улучшения прогноза, ε — среднеквадратичные ошибки аппроксимации.

Однако такой подход не может в действительности решить поставленную задачу, потому что при наличии опосредованного воздействия $Z \rightarrow Y \rightarrow X$ с его помощью будет обнаружена также связь $Z \rightarrow X$, так как X содержит информацию о динамике Z . А при воздействии $Z \rightarrow Y$ и $Z \rightarrow X$ может быть обнаружено воздействие $Y \rightarrow X$, поскольку Y может содержать часть информации от Z , что поможет предсказать X . Чтобы

Соответствие искомой архитектуры связей и выводов о наличии связи при идеальной чувствительности и специфичности предложенной методики (3), (4)

№ n/n	Вид взаимодействия	$PI_{z \rightarrow y}$	$PI_{y \rightarrow x z}$	$PI_{z \rightarrow x y}$
1	$Z \rightarrow Y \rightarrow X$	1	1	0
2	$Z \rightarrow X \wedge Z \rightarrow Y \rightarrow X$	1	1	1
3	$Z \rightarrow Y \wedge Z \rightarrow X$	1	0	1
4	$Z \rightarrow X \wedge Y \rightarrow X$	0	1	1
5	$Z \wedge Y \rightarrow X$	0	1	0
6	$Z \rightarrow X \wedge Y$	0	0	1
7	$Z \rightarrow Y \wedge X$	1	0	0
8	$Z \wedge Y \wedge X$	0	0	0

Примечание. Белый фон (отвечает $PI = 1$) соответствует значимым выводам о наличии связи, серый (отвечает $PI = 0$) — об отсутствии.

решить проблему, нужно вместо простой причинности по Грейнджеру, когда системы рассматриваются попарно, перейти к условной [2], построив модели вида

$$x'_{n+\tau|Y} = f_{xy}(\mathbf{x}_n, \mathbf{y}_n), \quad x'_{n+\tau|Z} = f_{xz}(\mathbf{x}_n, \mathbf{z}_n), \quad x'_{n+\tau|Y,Z} = f_{xyz}(\mathbf{x}_n, \mathbf{y}_n, \mathbf{z}_n),$$

$$y'_{n+\tau} = f_y(\mathbf{y}_n), \quad y'_{n+\tau|Z} = (y_n, \mathbf{z}_n) \quad (3)$$

и рассчитав улучшение прогноза по формулам

$$PI_{y \rightarrow x|z} = 1 - \frac{\varepsilon_{xyz}^2}{\varepsilon_{xz}^2}, \quad PI_{z \rightarrow x|y} = 1 - \frac{\varepsilon_{xyz}^2}{\varepsilon_{xy}^2}, \quad PI_{z \rightarrow y} = 1 - \frac{\varepsilon_{yz}^2}{\varepsilon_y^2}. \quad (4)$$

Результаты, которые должны получиться с использованием такой методики в „идеальной“ ситуации, показаны в таблице. Видно, что для каждого из восьми выбранных случаев возможного взаимодействия существует только одна комбинация результатов методики, т. е. достигается взаимно однозначное соответствие.

В реальной ситуации достижение $PI = 0$ (при отсутствии связей) и $PI = 1$ (при наличии) или даже близких к ним с точностью вычислений значений невозможно как по причине конечности частоты и объема выборки [7], так и из-за несовершенства моделей [8]. Поэтому принято

считать, что связь обнаружена, если полученное значение выше тестового уровня, для расчета которого в настоящей работе использовался широко распространенный подход, основанный на генерации суррогатных временных рядов [9]. Генерировалось 100 реализаций суррогатных рядов тестируемых систем с теми же параметрами при различных начальных условиях, но при отсутствии связи, по ним оценивался 95% уровень значимости. В [10] было показано, что такой способ дает удовлетворительные результаты и примерно эквивалентен некоторым другим распространенным подходам.

Размерность вектора состояния и степень полинома в моделях (3) выбирались по критерию Шварца [11]. Параметры (лаг вложения l и дальность прогноза τ) выбирались с учетом основного временного масштаба ведомой системы, как предложено в [12].

Для тестирования в качестве систем X , Y и Z были выбраны широко известные эталонные модели радиотехнических и оптических систем: уравнения генератора с 1.5 степенями свободы [13] и система Икеды [14]. Параметры систем и связи были выбраны и введены в соответствии с уравнениями

$$\begin{cases} \dot{x}_1 = (F(x_3, 26) - x_1)/3 + k_{yx}(y_1 - x_1) + k_{zx}(z_1 - x_1), \\ \dot{x}_2 = x_1 - x_3, \\ \dot{x}_3 = x_2 - 0.21x_3, \\ \dot{y}_1 = (F(y_3, 28) - y_1)/3 + k_{zy}(z_1 - y_1), \\ \dot{y}_2 = y_1 - y_3, \\ \dot{y}_3 = y_2 - 0.22y_3, \\ \dot{z}_1 = (F(z_3, 30) - z_1)/3, \\ \dot{z}_2 = z_1 - z_3, \\ \dot{z}_3 = z_2 - 0.23z_3, \end{cases}$$

$$F(x, M) = Mx \exp(x^2) \quad (5)$$

и

$$\begin{aligned} \dot{x}(t) &= -x(t) + 15 \sin(x(t-3) - x_0) + k_{yx}(y(t) - x(t)) + k_{zx}(z(t) - x(t)), \\ \dot{y}(t) &= -y(t) + 20 \sin(y(t-2) - y_0) + k_{zy}(z(t) - y(t)), \\ \dot{z}(t) &= -z(t) + 25 \sin(z(t-4) - z_0). \end{aligned} \quad (6)$$

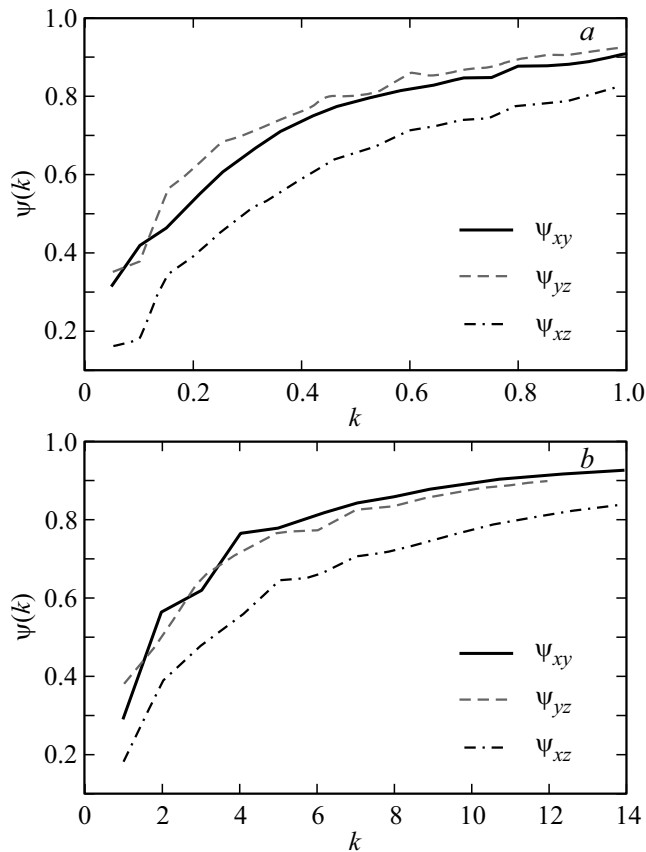


Рис. 1. Зависимости значений коэффициентов фазовой синхронизации ψ_{xy} , ψ_{xz} и ψ_{yz} от коэффициентов связи $k_{yx} = k_{zy} = k$: *a* — между автогенераторами с 1.5 степенями свободы в уравнении (5), *b* — между системами Икеды (6). Коэффициент $k_{zx} = 0$, что соответствовало архитектуре 1.

Для системы (5) координаты x_1 , y_1 , z_1 считались наблюдаемыми. Все или некоторые коэффициенты связи k_{yx} , k_{zx} и k_{zy} могли обнуляться, чтобы рассмотреть различные архитектуры связи. Если соответствующий коэффициент не обнулялся, он изменялся в широком диапазоне, как показано на рис. 1, *a*. Были выбраны три значения k : 0.15, 0.4 и 0.7,

для которых проводилось исследование с усреднением результата по 100 реализациям, полученным при различных начальных условиях.

Результаты для связанных автогенераторов представлены на рис. 2, *a*, где показано, что в среднем наилучшую работоспособность метод демонстрирует в случае, когда $k = 0.4$. Однако при наличии прямого и опосредованного взаимодействия ($Z \rightarrow X \wedge Z \rightarrow Y \rightarrow X$) наилучшие результаты достигаются при $k = 0.7$. Следует отметить, что наиболее сложным оказался тест для оценки $PI_{y \rightarrow x|z}$ в случае $Z \rightarrow Y \wedge Z \rightarrow X$, где наихудшие результаты достигаются при $k = 0.7$. По остальным возможным взаимодействиям метод демонстрирует ожидаемые результаты.

Аналогичным образом численный эксперимент проводился и для связанных однонаправленной связью систем Икеды (6). На основании зависимости, представленной на рис. 1, *b*, были выбраны три значения $k : 2, 5, 10$, для которых проводилось три теста с усреднением по 100 реализациям. Результаты численного эксперимента представлены на рис. 2, *b*. Метод демонстрирует наилучшую работоспособность для $k = 2$. При этом ошибочные выводы о наличии направленной связи наблюдаются для оценок $PI_{y \rightarrow x|z}$ и $PI_{z \rightarrow x|y}$ в случаях $Z \rightarrow Y \wedge Z \rightarrow X$ и $Z \rightarrow Y \rightarrow X$ соответственно. По остальным возможным взаимодействиям метод демонстрирует ожидаемые результаты.

В результате численного эксперимента было показано, что с помощью предложенной методики по скалярным временным рядам сложных колебаний трех систем с выраженным характерным временным масштабом можно надежно различить прямое и опосредованное воздействие в 82% случаев, используя модели, построенные с учетом временных масштабов сигналов. Эффективность метода выше для некоторых „средних“ коэффициентов связи, когда наличие связей заметно модифицирует динамику ведомой системы, но при этом не приводит к синхронизации. Полученный результат ценен тем, что оказывается возможным различать источник и путь распространения сигнала, не имея априорных знаний о системе, его генерирующей.

Наибольшую сложность представляет случай наличия одновременно прямого и опосредованного взаимодействия. Такой случай часто не может быть отделен от случая наличия только прямого воздействия первой системы в цепочке на две остальные: конечную и промежуточную. Это объясняется слабым вкладом опосредованного воздействия при наличии прямого в динамику третьей системы. Однако на практике ошибки такого рода, как правило, наименее критичны. Также показано,

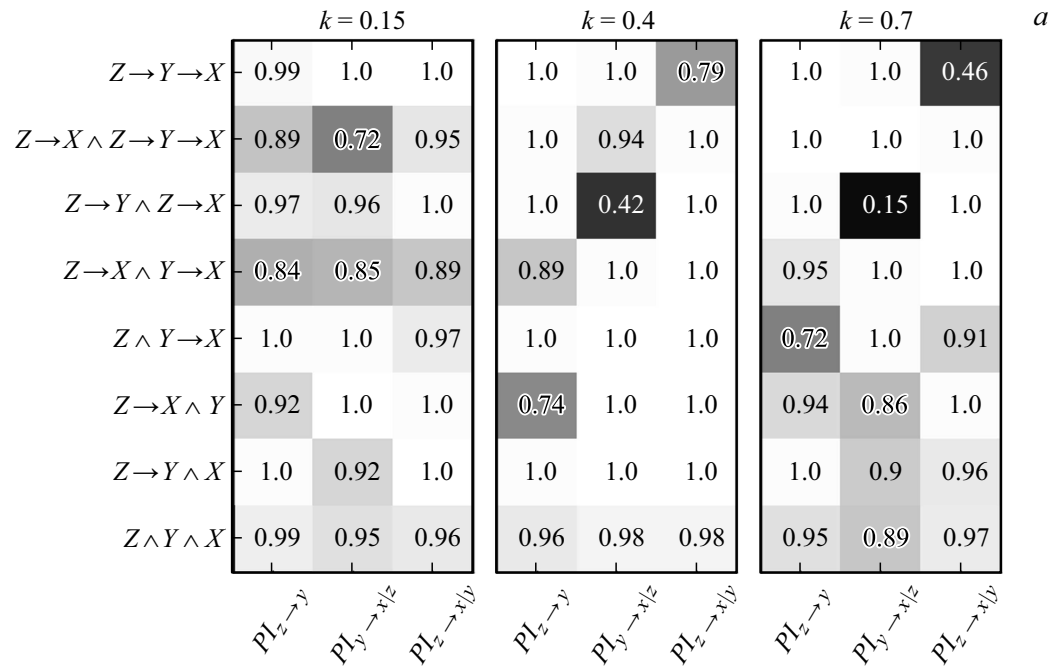


Рис. 2. Процент верных выводов о наличии (отсутствии) связи между системами по выбранным восьми видам взаимодействия для автогенераторов с 1.5 степенями свободы (*a*) и для систем Икеды (*b*). Для каждого типа прогностической модели грациями серого (белый — 100%, черный — 0%) и значениями отмечены отношения верных выводов (в соответствии с таблицей) к общему количеству тестов.

b

	$k = 2$			$k = 5$			$k = 10$		
$Z \rightarrow Y \rightarrow X$	1.0	1.0	0.87	1.0	1.0	0.03	1.0	1.0	0.0
$Z \rightarrow X \wedge Z \rightarrow Y \rightarrow X$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$Z \rightarrow Y \wedge Z \rightarrow X$	1.0	0.63	1.0	1.0	0.0	1.0	1.0	0.0	1.0
$Z \rightarrow X \wedge Y \rightarrow X$	0.92	1.0	1.0	0.92	1.0	1.0	0.99	1.0	1.0
$Z \wedge Y \rightarrow X$	0.98	1.0	0.93	0.97	1.0	0.82	0.96	1.0	0.97
$Z \rightarrow X \wedge Y$	0.98	0.94	1.0	0.9	0.93	1.0	0.95	0.86	1.0
$Z \rightarrow Y \wedge X$	1.0	0.93	0.93	1.0	0.71	0.79	1.0	0.88	0.83
$Z \wedge Y \wedge X$	0.98	0.97	0.9	0.96	0.93	1.0	0.98	0.94	0.99
	$PI_{z \rightarrow y}$	$PI_{y \rightarrow x/z}$	$PI_{z \rightarrow x/y}$	$PI_{z \rightarrow y}$	$PI_{y \rightarrow x/z}$	$PI_{z \rightarrow x/y}$	$PI_{z \rightarrow y}$	$PI_{y \rightarrow x/z}$	$PI_{z \rightarrow x/y}$

Рис. 2 (продолжение).

что использованный подход пригоден для выявления факта отсутствия связей в цепочке или ее разрыва, поскольку он позволяет определить отсутствующие связи. При таком использовании его эффективность достигает 90% и более.

Результаты, полученные при исследовании цепочек из двух различных типов эталонных осцилляторов, в основном сходны, хотя для систем автогенераторов с 1.5 степенями свободы эффективность метода в среднем выше, что может объясняться более простой динамикой отдельных систем (меньшая размерность) и более выраженным временным масштабом колебаний. Это позволяет лучше подобрать параметры метода на основании ранее разработанных в рамках проекта критериев.

Работа выполнена при финансовой поддержке Российского научного фонда грант (14-12-00291).

Список литературы

- [1] *Granger C.W.J.* // *Econometrica*. 1969. V. 37. N 3. P. 424–438.
- [2] *Chen Y., Rangarajan G., Feng J., Ding M.* // *Phys. Lett. A*. 2004. V. 324. N 1. P. 26–35.
- [3] *Smirnov D.A., Bezruchko B.P.* // *Europhys. Lett.* 2012. V. 100. N 1. P. 10005.
- [4] *Angelini L., Pellicoro M., Stramaglia S.* // *Phys. Lett. A*. 2009. V. 373. N 29. P. 2467–2470.
- [5] *Yick J., Mukherjee B., Ghosal D.* // *Comput. Networks*. 2008. V. 52. N 12. P. 2292–2330.
- [6] *Packard N., Crutchfield J., Farmer J., Shaw R.* // *Phys. Rev. Lett.* 1980. V. 45. N 9. P. 712–716.
- [7] *Крылов С.Н., Смирнов Д.А., Осипов Г.В., Безручко Б.П.* // *Письма в ЖТФ*. 2015. Т. 41. В. 11. С. 94–102.
- [8] *Smirnov D.A.* // *Phys. Rev. E*. 2014. V. 90. N 6. P. 062921.
- [9] *Theiler J., Eubank S., Longtin A., Galdrikian B., Farmer J.D.* // *Physica D*. 1992. V. 58. N 1-4. P. 77–94.
- [10] *Корнилов М.В., Сысоев И.В.* // *Изв. вузов. Прикладная нелинейная динамика*. 2014. Т. 22. № 4. С. 66–75.
- [11] *Schwarz G.* // *Annals Statist.* 1978. V. 6. N 2. P. 461–464.
- [12] *Kornilov M.V., Medvedeva T.M., Bezruchko B.P., Sysoev I.V.* // *Chaos, Solit. Fract.* 2016. V. 82. P. 11–21.
- [13] *Дмитриев А.С., Кислов В.Я.* // *Радиотехника и электроника*. 1984. Т. 29. № 12. С. 2389–2398.
- [14] *Ikeda K., Matsumoto K.* // *Physica D*. 1987. V. 29. N 1-2. P. 223–235.