

Технічні науки

УДК 004.021

Хомицевич Микита Володимирович

студент

*Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»*

Хомицевич Нитита Васильевич

студент

*Национального технического университета Украины
«Киевский политехнический институт имени Игоря Сикорского»*

Khomitsevich Nikita

Student of the

*National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute»*

**АРХІТЕКТУРА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ
ПРОВЕДЕННЯ ПРИЧИННО-НАСЛІДКОВОГО АНАЛІЗУ
АРХИТЕКТУРА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ
ПРОВЕДЕНИЯ ПРИЧИННО-СЛЕДСТВЕННОГО АНАЛИЗА
SOFTWARE ARCHITECTURE FOR THE CAUSE-EFFECT ANALYSIS**

Анотація. В роботі розглянуто поетапний алгоритм розробки програмного забезпечення для проведення причинно-наслідкового аналізу динамічних процесів представлених дискретними моделями.

Ключові слова: причинно-наслідковий аналіз, матриця причинно-наслідкових зв'язків, причинно-наслідковий зв'язок.

Аннотация. В работе рассмотрены поэтапный алгоритм разработки программного обеспечения для проведения причинно-

следственного анализа динамических процессов представленных дискретными моделями.

Ключевые слова: причинно-следственный анализ, матрица причинно-следственных связей, причинно-следственная связь.

Summary. The paper considers a step-by-step algorithm for software development for conducting a causal analysis of the dynamic processes presented by discrete models.

Key words: causation analysis, matrix of causal relationships, causal relationship.

Даний програмний продукт розраховано для автоматизування процесу причинно-наслідкового аналізу динамічних процесів представлених дискретними моделями. Ці дискретні моделі можуть бути представлені будь-якими даними. Важливою умовою для початку роботи з даним програмним забезпеченням є необхідність у формуванні Excel файлу з необхідними для аналізу даними. Далі будемо розглядати весь функціонал додатку на прикладі даних економіки України за 1992-1994 рр. [1].

Після відкриття програми користувач бачить головне вікно (рис. 1).

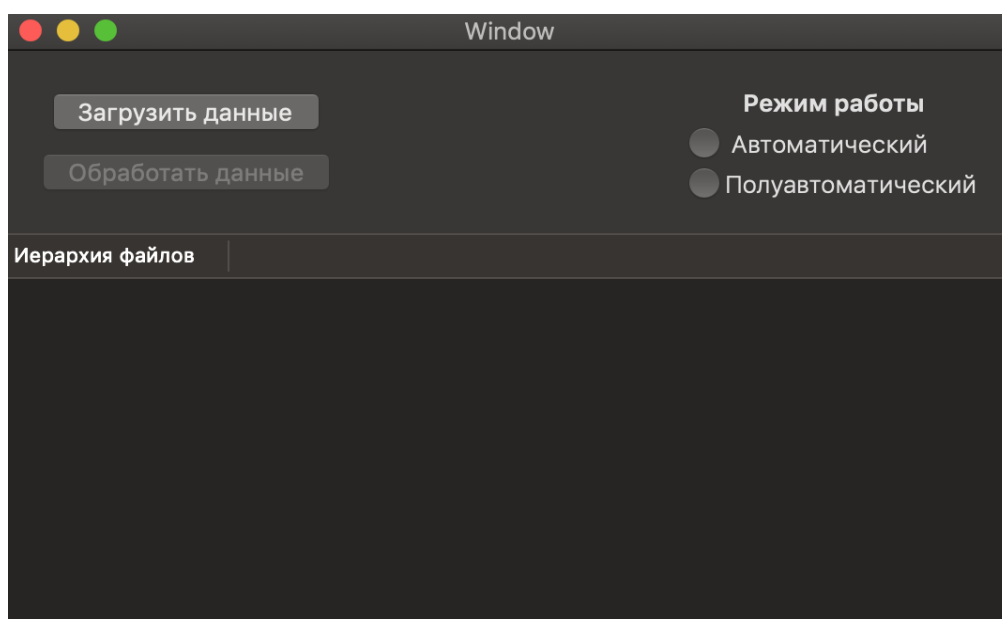


Рис. 1. Головне вікно програми

В цьому вікні користувач має всього декілька компонентів інтерфейсу (табл. 1), що надають функціональність тільки для завантаження даних у програму, вибір режиму роботи та початок обробки даних алгоритмами причино-наслідкового аналізу (порахувати матрицю причино-наслідкових зв'язків, визначити набір зовнішніх (екзогенних) змінних та побудувати граф зв'язків вхідних процесів заданої системи дискретних моделей).

Таблиця 1

Детальний опис елементів інтерфейсу першого екрану програми

Назва елемента	Його функція
UIButton "Загрузить данные"	Завантаження даних та відображення у таблиці
UIButton "Обработать данные"	Запуск роботи алгоритмів ПНА
UIRadioGroupButton «Автоматический» та «Полуавтоматический»	Один з двох можливих режимів роботи програми
UITableView «Иерархия файлов»	Візуальне відображення набору імпортованих файлів згідно назв

Перше, що має зробити користувач – це завантажити збережений із попередньої програми МАКСО набір даних, але є певний ряд обмежень. По-перше, кожен Excel файл має бути названий відповідно до того, які пари зміни він описую. Приклад: x1 – вихідна змінна, x2 – вхідна змінна. Тоді назва файлу має бути наступною – "x01_x01.xlsx". Таким чином ми одразу описуємо який файл за яку пару з системи моделей прогнозу буде відповідати. По-друге, усі файли мають бути згруповані між собою в набір каталогів, кожен з яких також має називатися певним чином. Приклад: якщо вихідна змінна для певного набору Excel файлів – x1, то каталог має бути названий так само, як і назва цієї змінної. Це також зроблено для більш простого процесу аналізу вхідних даних.

Коли усі файли були згруповані згідно з вищевказаними правилами, користувач може завантажити усі дані в програму. Для цього він має натиснути на кнопку «Загрузить данные». Після натиску на кнопку

з'явиться нове вікно, в якому користувач має вказати шлях до наших даних. Дані Це є максимально просто, бо користувачу надається візуальне представлення усіх його директорій, що є в нього на диску (рис. 2).

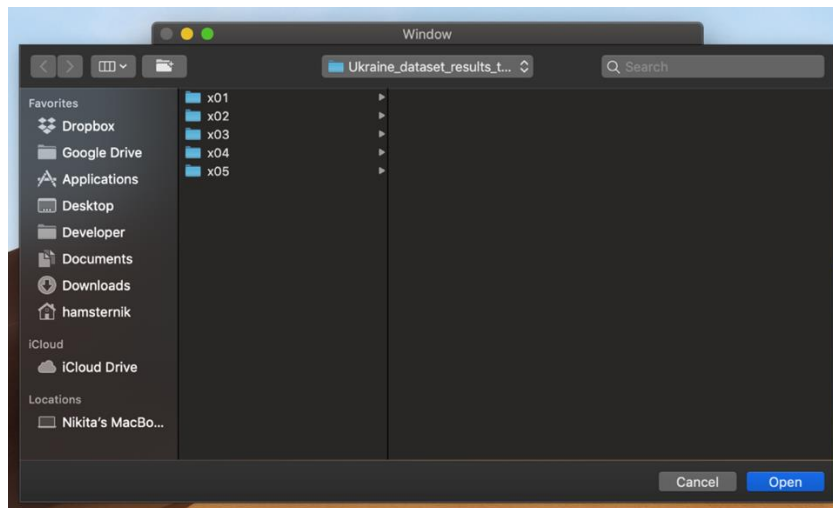


Рис. 2. Візуальне представлення директорій диска користувача

Коли користувач вказав вірний шлях до потрібних програмі даних, на головному вікні програми в нижній половині відображається набір завантажених змінних, кожна з яких має власний набір вхідних змінних (рис. 3). Усі дані були отримані за рахунок аналізу та парсингу завантажених Excel файлів.

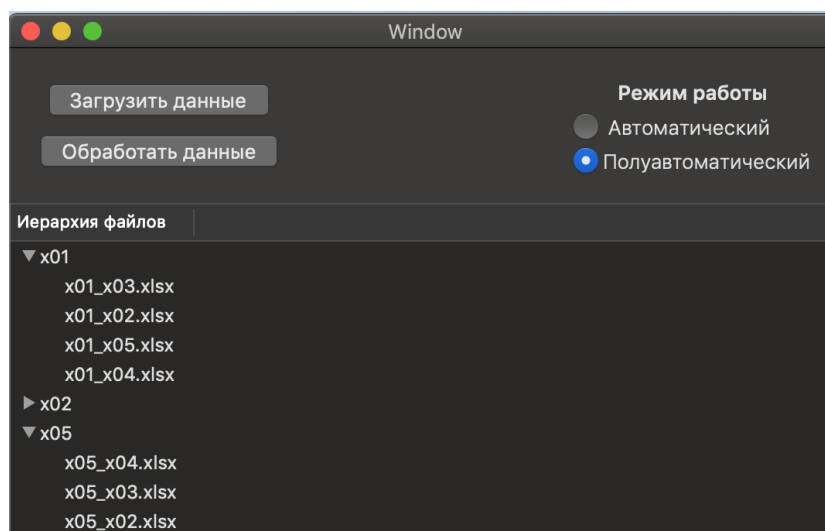


Рис. 3. Головне вікно із завантаженими даними

Після завантаження даних для користувача стає активною до натискання друга кнопка на екрані «Обработать данные». Саме після її натискання запуститься механізм обробки даних заданим набором алгоритмів для проведення причино-наслідкового аналізу. Але перед тим користувач має спочатку обрати необхідний для нього режим роботи.

Перейдемо до подальшого опису функціонування різних режимів додатку. Всього в програмі присутні 2 режими:

1. Напівавтоматичний режим

В цьому режимі користувач спочатку отримає нове вікно програми (рис. 4), в якому йому відображаються формули прямої та оберненої моделей відповідно до вхідних/вихідних змінних, значення похибки дисперсії кореляції зв'язку $\delta_{x_i}(x_j)$ та $\delta_{x_j}(x_i)$ а також підсумування – результат алгоритму порівняння похибок дисперсій отриманих моделей [2].

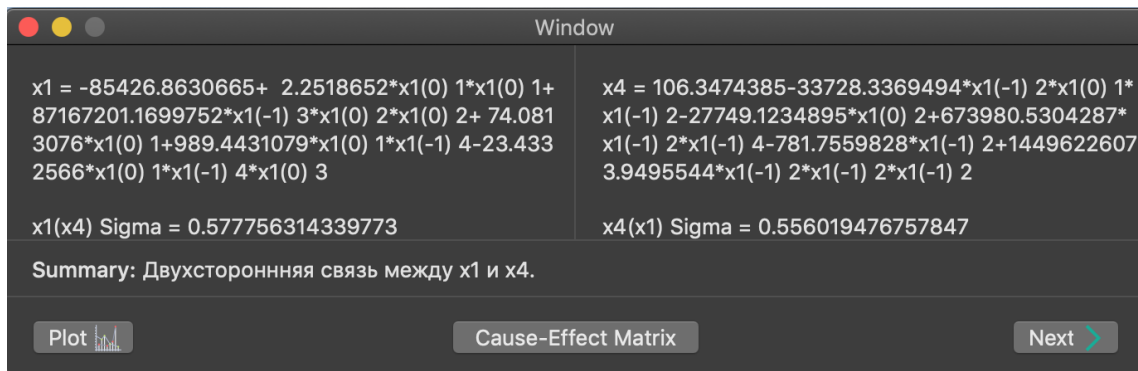


Рис. 4. Вікно відображення конкуруючих моделей прогнозу

Далі користувач може або натиснути кнопку “Next” та дивитись ізольовані пари прямої та оберненої моделей для усього набору параметрів системи, або кнопку “Cause-Effect Matrix” та перейти на новий екран з результатами у вигляді матриці причино-наслідкових зв'язків.

Є п'ять різних варіантів тексту для даного елемента (табл. 2), який залежить від результату алгоритму порівняння помилок між прямою та оберненою моделями.

Текстове представлення результату алгоритму порівняння

Варіант	Текст варіанту
$x_i = -1, x_j = -1$	Взаимосвязь между x_i и x_j не установлена.
$x_i = 1, x_j = 0$	x_i : следствие параметра x_j .
$x_i = 0, x_j = 1$	x_i : причина параметра x_j .
$x_i = 1, x_j = 1$	Двухсторонняя связь между x_i и x_j .
$x_i = 0, x_j = 0$	Переменные x_i и x_j не взаимосвязаны.

2. Автоматичний режим

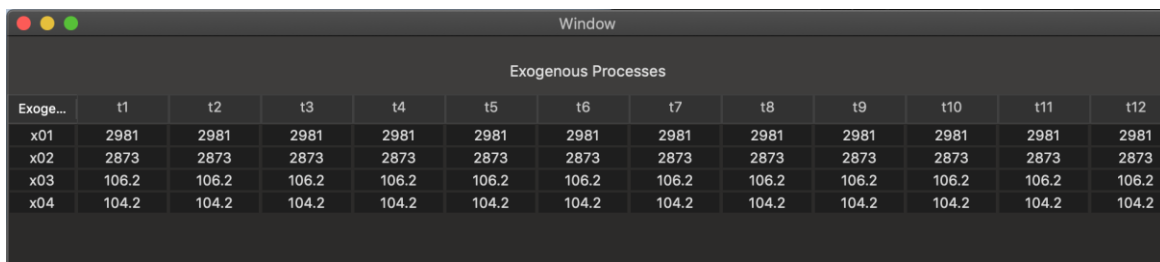
В цьому режимі користувач після натиску кнопки «Обработать данные» одразу отримає вікно з результатами (рис. 5), а саме матрицею причино-наслідкових зв'язків. Цей режим потрібен коли користувач впевнений в тих даних, які він завантажив в програму і коли він власноруч бачив ті існуючі пари дискретних моделей системи процесів. Йому не потрібно дивитись на похибки кореляції тощо.

PTitles	x00	x01	x02	x03	x04
x01	-	1	1	1	1
x02	1	-	1	1	1
x03	1	1	-	1	1
x04	1	1	1	-	1
x05	1	1	1	0	-

Рис. 5. Вікно з результатами роботи програми

На цьому екрані користувач може вже подивитись перші отримані результати програми. Саме на базі матриці причино-наслідкових зв'язків і буде проводитись подальший аналіз та виконувати роботу запрограмовані алгоритми [3].

Користувач може подивитись набір екзогенних (зовнішніх) змінних системи, що ніяким чином не впливають на встановлення статистичної причино-наслідкової структури динамічних характеристик системи процесів і які можна буде в подальшому аналізі виключити. Для цього користувач повинен натиснути на верхню кнопку "Show" (що під назвою «Exogenous parameters»). За декілька секунд програма порахує набір зовнішніх змінних та відкриє нове вікно (рис. 6), де буде наглядно відображена множина цих параметрів.



Exoge...	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12
x01	2981	2981	2981	2981	2981	2981	2981	2981	2981	2981	2981	2981
x02	2873	2873	2873	2873	2873	2873	2873	2873	2873	2873	2873	2873
x03	106.2	106.2	106.2	106.2	106.2	106.2	106.2	106.2	106.2	106.2	106.2	106.2
x04	104.2	104.2	104.2	104.2	104.2	104.2	104.2	104.2	104.2	104.2	104.2	104.2

Рис. 6. Вікно множини екзогенних змінних та їх значень

Далі користувач може подивитись граф залежностей [4], який встановлює залежність змінних системи, що ніяким чином не впливають на встановлення статистичної причино-наслідкової структури динамічних характеристик системи. Для цього він має натиснути на кнопку "Show" з назвою "Dependency graph". В результаті користувач може побачити наступне зображення (рис. 7).

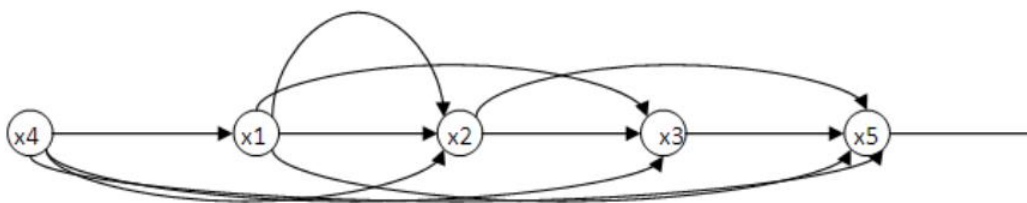


Рис. 7. Вікно множини екзогенних змінних та їх значень

Таким чином ми бачимо, що в нас параметр $x4$ є найбільш впливовою серед усіх інших. І саме це приводить до останнього алгоритму який демонструє які нові статистичні моделі на основі проведеного причино-наслідкового аналізу можна отримати. Маючи наступний граф ми можемо

отримати нову модель з декількома вхідними змінними для кожної вихідної змінної з урахуванням податкових результатів моделей.

Для цього на вікні результатів роботи програми внизу ми можемо бачити кнопку з іконкою “Search” та поверх текстовим полем для введення назви змінної.

Таким чином було розроблене програмне забезпечення, що надає користувачу можливість проведення причинно-наслідкового аналізу. В результаті користувач може побудувати матрицю причинно-наслідкових зв'язків а також використати набір пропонованих алгоритмів для подальшого проведення причинно-наслідкового аналізу, а саме: визначення множини зовнішніх змінних, побудова причинно-наслідкового фільтру та побудова графа причинно-наслідкової структури [5].

Література

1. M. A. T. Figueiredo and A. K. Jain. Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3): 381–396, March 2002.
2. F. Eberhardt and R. Scheines. Interventions and causal inference. *Philosophy of Science*, 74(5): 981–995, 2007.
3. P. Ein-Dor and J. Feldmesser. Attributes of the performance of central processing units: a relative performance prediction model. *Communications of the ACM*, 30:308–317, 1987.
4. R. M. Gray. Toeplitz and circulant matrices: A review. *Foundations and Trends in Communications and Information Theory*, 2:155–239, 2006.
5. Alexander S. New methods for separating causes from effects in genomics data [Електронний ресурс] / S. Alexander, H. Mikael, L. Nikita – 2012. – Режим доступу:
<https://bmcgenimcs.biomedcentral.com/articles/10.1186/1471-2164-13-S8-S22>.