

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

СОГЛАСОВАНО

Генеральный директор
ЗАО «АйТи»



Бакуев О.Р.

2011 г.

УТВЕРЖДАЮ

Ректор НИУ ИТМО



Васильев В.Н.

2011 г.

МНОГОПРОФИЛЬНАЯ ИНСТРУМЕНТАЛЬНО-
ТЕХНОЛОГИЧЕСКАЯ ПЛАТФОРМА СОЗДАНИЯ
И УПРАВЛЕНИЯ РАСПРЕДЕЛЕННОЙ СРЕДОЙ
ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ CLAVIRE

ПРИКЛАДНОЙ СЕРВИС РАСЧЕТА ЭФФЕКТИВНОСТИ СТРАТЕГИИ
ВАКЦИНАЦИИ ДЛЯ ПРОТИВОДЕЙСТВИЯ РАЗВИТИЮ ВИРУСНОЙ
ЭПИДЕМИИ

ОПИСАНИЕ ПРОГРАММЫ

ЛИСТ УТВЕРЖДЕНИЯ

RU.СНАБ.80066-06 13 51-ЛУ

Представители
Организации-разработчика

Руководитель разработки,
профессор НИУ ИТМО

Бухановский А.В.
"29" декабря 2011 г.

Ответственный исполнитель,
с.н.с. НИУ ИТМО

Луценко А.Е.
"29" декабря 2011 г.

Нормоконтролер
ведущий инженер НИУ ИТМО

Позднякова Л.Г.
"29" декабря 2011 г.

2011

Ине.№ подл.	Подп. и дата
Взам.ине.№	Подп. и дата
Ине.№ дубл.	Подп. и дата

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**УТВЕРЖДЕН
RU.СНАБ.80066-06 13 51-ЛУ**

**МНОГОПРОФИЛЬНАЯ ИНСТРУМЕНТАЛЬНО-
ТЕХНОЛОГИЧЕСКАЯ ПЛАТФОРМА СОЗДАНИЯ
И УПРАВЛЕНИЯ РАСПРЕДЕЛЕННОЙ СРЕДОЙ
ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ CLAVIRE**

**ПРИКЛАДНОЙ СЕРВИС РАСЧЕТА ЭФФЕКТИВНОСТИ СТРАТЕГИИ
ВАКЦИНАЦИИ ДЛЯ ПРОТИВОДЕЙСТВИЯ РАЗВИТИЮ ВИРУСНОЙ
ЭПИДЕМИИ**

ОПИСАНИЕ ПРОГРАММЫ

RU.СНАБ.80066-06 13 51

ЛИСТОВ 28

Инв.№ подл.		Подп. и дата	
Взам. инв. №		Инв. № дубл.	

АННОТАЦИЯ

Документ содержит описание прикладного сервиса расчета эффективности стратегии вакцинации для противодействия развитию вирусной эпидемии. Прикладной сервис реализует композитное приложение для многоцелевой инструментально-технологической платформы (МИТП) CLAVIRE, которое позволяет, основываясь на заданных вероятностных характеристиках топологии контактной эпидемиологической сети, рассчитать интервальные оценки динамики количества зараженных индивидов в популяции, исходя из различных стратегий вакцинации. Прикладной сервис разработан в ходе выполнения проекта «Создание распределенной вычислительной среды на базе облачной архитектуры для построения и эксплуатации высокопроизводительных композитных приложений» (Договор № 21057 от 15 июля 2010 г., шифр 2010-218-01-209) в рамках реализации постановления Правительства РФ № 218 «О мерах государственной поддержки развития кооперации российских высших учебных заведений и организаций, реализующих комплексные проекты по созданию высокотехнологичного производства».

СОДЕРЖАНИЕ

1 ОБЩИЕ СВЕДЕНИЯ	4
2 ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ	5
2.1 Область применения	5
2.2 Функциональное назначение.....	5
2.3 Ограничения на применение	6
3 ОПИСАНИЕ ЛОГИЧЕСКОЙ СТРУКТУРЫ.....	6
3.1 Структура композитного приложения	6
3.2 Характеристики модулей композитного приложения	9
3.3 Описание модулей композитного приложения в МИТП	14
4 ИСПОЛЬЗУЕМЫЕ ТЕХНИЧЕСКИЕ СРЕДСТВА	20
5 ВЫЗОВ И ЗАГРУЗКА.....	21
6 ВХОДНЫЕ ДАННЫЕ	22
7 ВЫХОДНЫЕ ДАННЫЕ	23
ПЕРЕЧЕНЬ СОКРАЩЕНИЙ	26
ПЕРЕЧЕНЬ ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	27

1. ОБЩИЕ СВЕДЕНИЯ

Прикладной сервис (ПС) расчета эффективности стратегии вакцинации для противодействия развитию вирусной эпидемии RU.СНАБ.80066-06 01 51 реализует композитное приложение, которое позволяет, основываясь на заданных вероятностных характеристиках топологии контактной эпидемиологической сети, рассчитать интервальные оценки динамики количества зараженных индивидов в популяции, исходя из различных стратегий вакцинации. ПС разработан в ходе выполнения проекта «Создание распределенной вычислительной среды на базе облачной архитектуры для построения и эксплуатации высокопроизводительных композитных приложений» (Договор № 21057 от 15 июля 2010 г., шифр 2010-218-01-209) в рамках реализации постановления Правительства РФ № 218 «О мерах государственной поддержки развития кооперации российских высших учебных заведений и организаций, реализующих комплексные проекты по созданию высокотехнологичного производства».

ПС функционирует в рамках распределенной среды облачных вычислений под управлением многофункциональной инструментально-технологической платформы (МИТП) CLAVIRE RU.СНАБ.80066-06. Он разработан на предметно-ориентированном языке EasyFlow описания композитных приложений.

ПС использует следующие пакеты прикладных программ, доступные в распределенной среде под управлением МИТП (см. также раздел 3.2):

- CNM - прикладной пакет генерации комплексной сети по заданному распределению топологических характеристик;
- VaccinStrategy - прикладной пакет цензурирования комплексной сети в соответствии с заданной стратегией;
- ISM - прикладной пакет распространения информации по сети с заданной топологией;
- SciLab - многоцелевой пакет компьютерной математики и визуализации (заимствуется).

Перечисленные пакеты описываются на языке EasyPackage и регистрируются в базе пакетов МИТП.

2. ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ

2.1. Область применения

ПС предназначен для моделирования и изучения процессов распространения вирусной эпидемии на основе аппарата комплексных сетей [1]. Его основным достоинством является возможность задавать им структуру общественных связей в *контактной* сети (различной для разных заболеваний). Например, для эпидемии гриппа рассматривается сеть контактов при перемещениях индивидов в местах общего пользования, а для ВИЧ и гепатита-С – сеть сексуальных контактов [2]. Распространение инфекции в контактной сети характеризуется динамикой взаимодействия узлов (индивидов), находящихся в одном из трех основных состояний: восприимчивый к вирусу, зараженный, удаленный из сети в результате обнаружения инфекции или смерти, или приобретенного иммунитета. Связи указывают на наличие контактов между людьми, которые могут приводить к новым случаям заражения. При этом в ходе эволюции сети связи могут изменяться как за счет внешних факторов (демография, смертность и пр.), так и за счет внутренних факторов, обусловленных сменой партнеров [3]. Для противодействия развитию эпидемии используются различные стратегии *вакцинации*, которые в рамках модели комплексной сети сводятся к различным схемам цензурирования (т.е. удаления ряда узлов и (или) ребер в соответствии с задаваемыми правилами). ПС позволяет, исходя из знаний о вероятностных характеристиках структуры сети, моделировать и исследовать эффективность различных стратегий вакцинации в целях обеспечения поддержки принятия решений при планировании соответствующих санитарно-эпидемиологических мероприятий.

2.2. Функциональное назначение

ПС предназначен для решения следующих задач:

- Моделирование неоднородной контактной сети заданного объема на основе распределений степеней узлов и характеристик связности между отдельными группами в рамках общей популяции
- Моделирование процесса распространения инфекции по контактной сети, исходя из заданной вероятности заражения и вероятности перехода в острую стадию (прекращения распространения инфекции данным узлом).

- Расчет интервальных характеристик (минимальные и максимальные порядковые статистики, медиана, размах выборки) при заданной стратегии вакцинации с учетом неопределенности и неполноты ее выполнения на сети заданного объема.

2.3. Ограничения на применение

ПС применим для моделирования распространения заболеваний, характеризуемых однородным течением латентной стадии (например, грипп), в предположении, что переход в острую стадию приводит к однозначной изоляции индивида. Для заболеваний, характеризуемых множественными рецессиями, в ходе которых индивид может вновь распространять инфекцию (например, ВИЧ), использование ПС ограничено первым приближением.

3. ОПИСАНИЕ ЛОГИЧЕСКОЙ СТРУКТУРЫ

3.1. Структура композитного приложения

Композитное приложение, реализуемое ПС, предназначено для моделирования выборки состояний популяции при развитии эпидемии, соответствующей различным вариантам вакцинации в рамках заданной стратегии (например, удаления заданной доли узлов случайным образом, или в соответствии с определенными признаками). Для этого необходимо по вероятностным характеристикам контактной сети сгенерировать модельную сеть, на основе которой исследуется распространение инфекции в условиях различных стратегий вакцинации. Для того, чтобы оценить неопределенность при выборе стратегии вакцинации, в простейшем случае используется *бутстрепирование* – формируется выборка на основе текущей сети с удалением части случайно выбранных узлов. Далее по каждому элементу выборки осуществляется моделирование распространения инфекции, чтобы в итоге определить интервальную оценку числа зараженных в заданный момент времени. Структура композитного приложения показана на рис. 3.1.

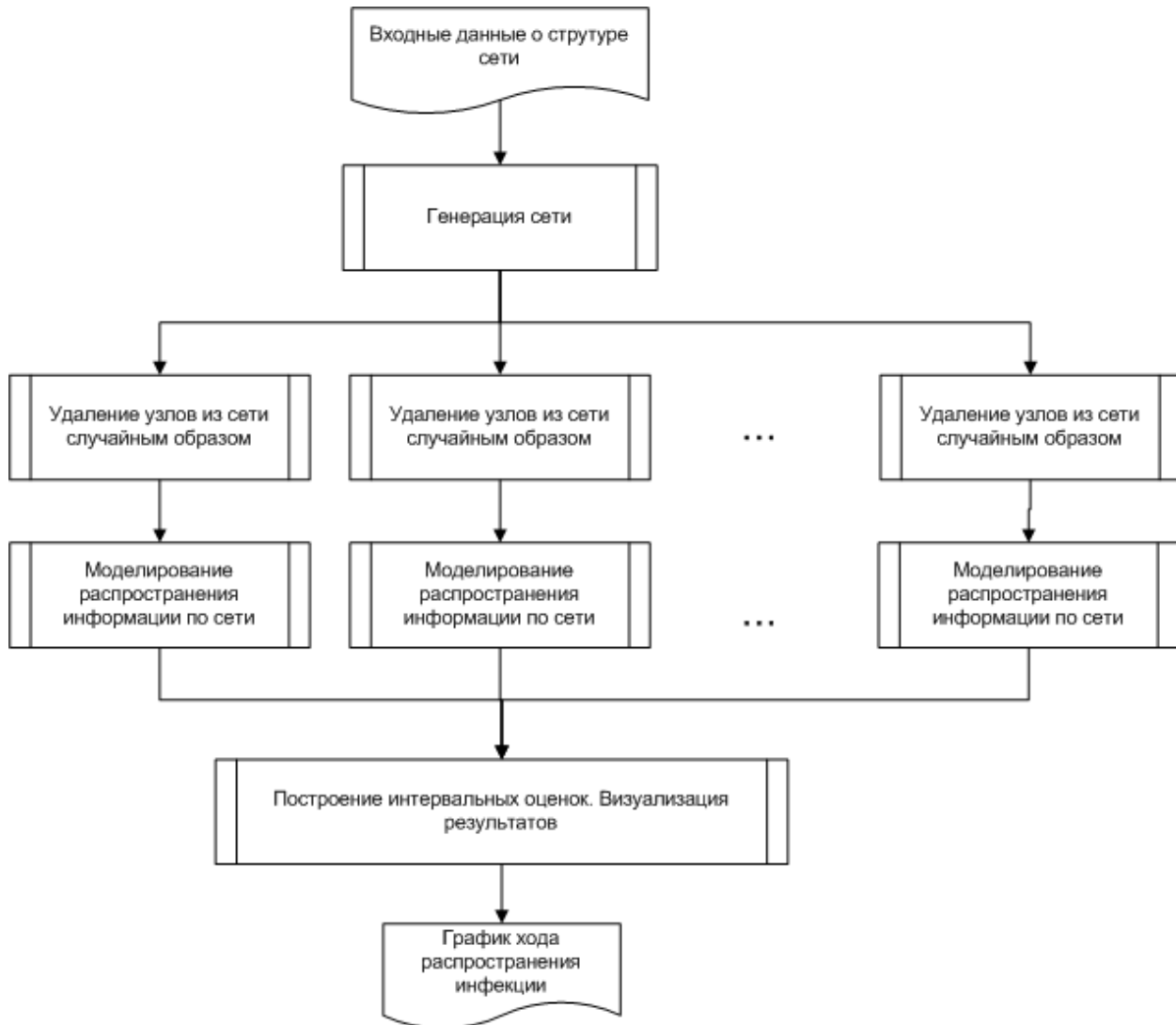


Рис. 3.1 – Структура композитного приложения

Для реализации композитного приложения в соответствии со структурой на рис. 3.1, используются следующие прикладные пакеты:

- CNM - прикладной пакет генерации комплексной сети по заданному распределению топологических характеристик;
- VaccinStrategy - прикладной пакет цензурирования комплексной сети в соответствии с заданной стратегией;
- ISM - прикладной пакет распространения информации по сети с заданной топологией;
- SciLab - многоцелевой пакет компьютерной математики и визуализации – необходим для построения интервальных оценок и визуализации результатов.

В листинге 3.1 приведен скрипт композитного приложения на языке EasyFlow, а на рис. 3.2 - результат его интерпретации в МИТП, который создает очередь из четырех последовательных процессов запуска пакетов в виде AWF.

Листинг 3.1. Описание композитного приложения на языке EasyFlow

```
require inNetwork, vaccinConfFile;
require infectionConfFile;

//Generate network
require inNetwork, vaccinConfFile;
require infectionConfFile;

//Generate network
step NetworkGeneration runs cnm (
    in_format = "short",
    inDataFile = inNetwork
);

//Exclude several nodes
step VaccinationStrategy runs VaccinStrategy after NetworkGeneration (
    inDataFile = sweep [NetworkGeneration.Result.outs["output.dat"],
                        NetworkGeneration.Result.outs["output.dat"],
                        NetworkGeneration.Result.outs["output.dat"],
                        NetworkGeneration.Result.outs["output.dat"]],
    inConfigFile = vaccinConfFile
);

//Infection spread Modelling
step InfectionSpreadModelling runs ism after VaccinationStrategy (
    inDataFile = sweep VaccinationStrategy.Result.sweep_outs["output.dat"],
    inConfigFile = infectionConfFile
);

//Result visualization
step Visualization runs scilab after InfectionSpreadModelling(
    script_name = "VSM_chart.sce",
    input_folder = InfectionSpreadModelling.Result.sweep_outs["output.dat"],
    output_file = "chart"
);
```

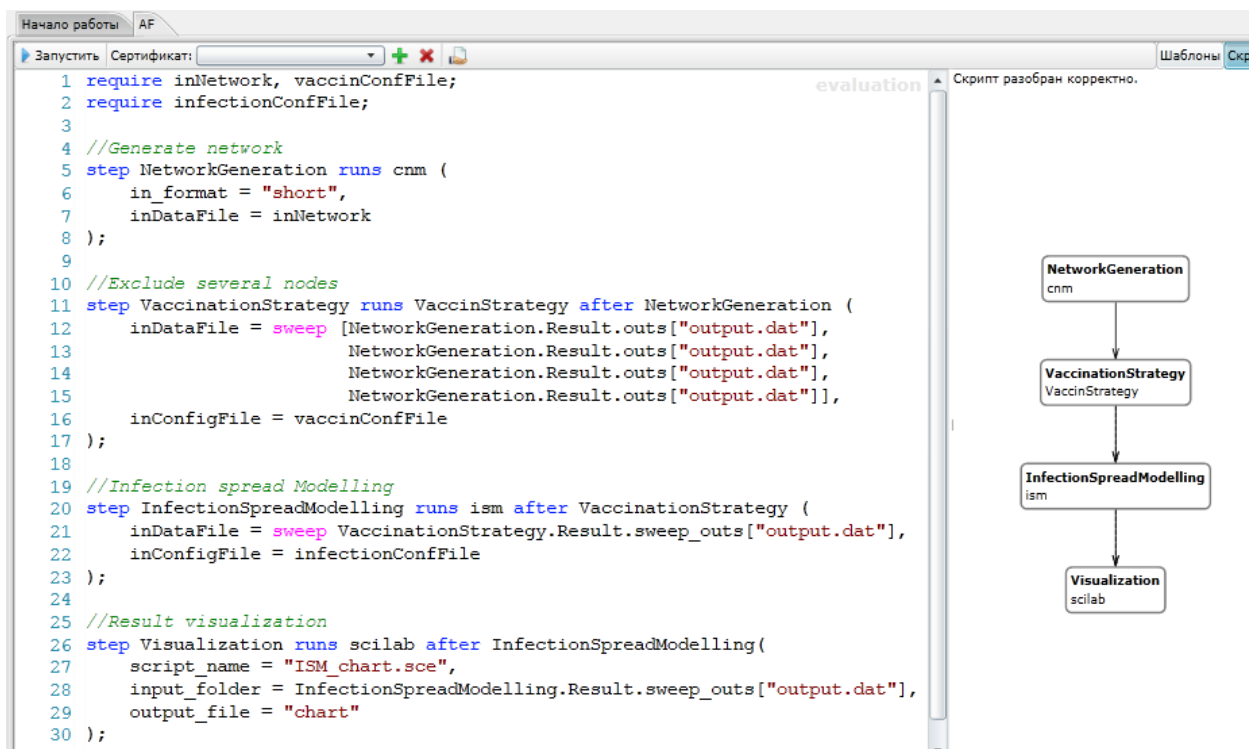


Рис. 3.2 – Подготовка композитного приложения к запуску на выполнение в МИТП

3.2. Характеристики модулей композитного приложения

3.2.1. Прикладной пакет CNM генерации комплексной сети

Прикладной пакет CNM – программный компонент моделирования комплексной неоднородной сети, позволяющий генерировать неоднородную сеть с заданными вероятностными характеристиками для последующего исследования различных процессов на ней. Пакет CNM разработан в виде консольного приложения с использованием языка программирования Java версии 1.6. CNM может функционировать на всех операционных системах, на которых может быть установлена платформа Java JRE 1.6.

CNM дает возможность методом Монте-Карло смоделировать комплексную сеть заданного объема, состоящую из нескольких классов вершин, где каждый класс вершин характеризуется собственным распределением степеней. При этом учитывается характер связей между узлами внутри каждого из классов, так и между ними. Результатом работы пакета является графовая структура, представимая в текстовом формате, применимом для дальнейшего моделирования.

В пакете CNM реализована расширенная конфигурационная модель комплексной сети. Для задания модели необходимо знать следующие параметры: Q – число классов

вершин (соответствующих подгруппам), $q \in \{1, 2, \dots, Q\}$ - индекс класса, n_q - число вершин класса q , d_i^q - степень вершины i класса q , D_q - степенная последовательность класса q , P_q - функция распределения степеней класса q , \overline{p}_q - средняя степень вершины класса q . Дополнительно должна быть задана $E = \{e_{ij}\}_{i=1..Q}^{j=1..Q}$ - матрица инцидентности классов размерности $q \times q$. Элемент e_{ij} ($0 \leq e_{ij} \leq 1$) показывает, какая доля исходящих ребер из вершин класса i связана с вершинами класса j . Алгоритм генерации сети является модификацией алгоритма Хавела-Хакими [4,5] для построения конфигурационной модели. Данный алгоритм позволяет построить сеть без петель и кратных ребер, что особенно актуально при моделировании социальных сетей, в которых они, как правило, не имеют места.

В качестве иллюстрации результатов расчетов пакетом CNM на рис. 3.2(а) изображена сеть, состоящая из двух классов вершин. Первый класс характеризуется степенным распределением с показателем $g = 1.5$, второй – Пуассоновым распределением со средним значением $l = 25$. На рисунке 3.3(б) изображена сеть, созданная по степенному распределению с $g = 2.5$.

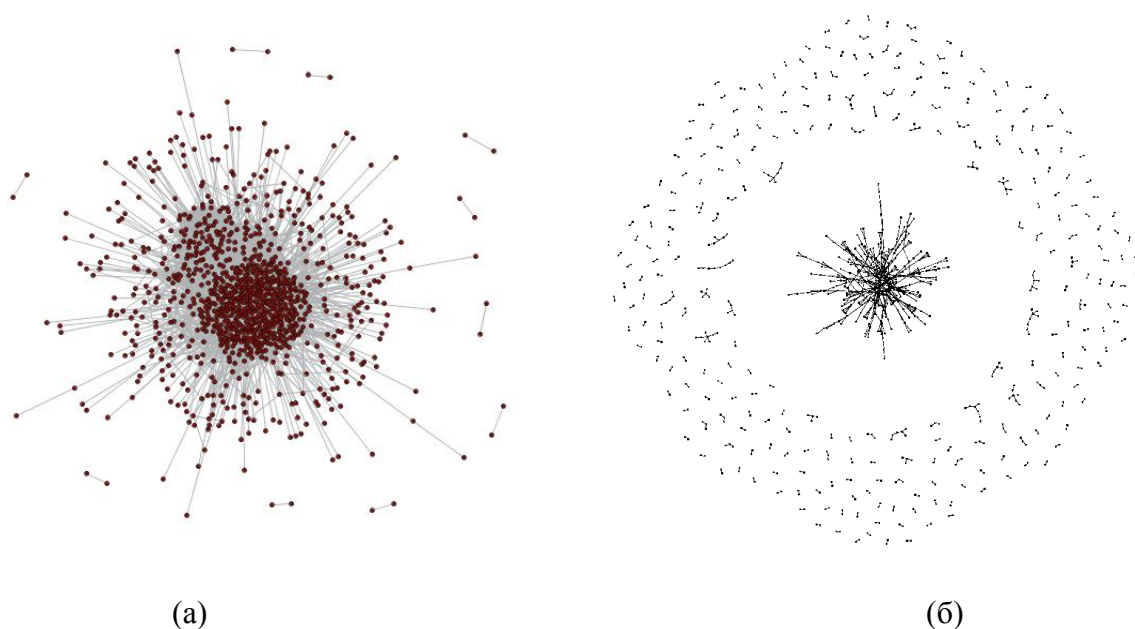


Рис. 3.3 – Примеры модельных комплексных сетей на основе расширенной конфигурационной модели (комментарии по тексту)

Приложение состоит из нескольких пакетов и классов, каждый из которых выполняет специфические функции. Ниже приведены наименования классов, и описание их функций.

`ru.ifmo.hpc.structure.Edge` — класс, описывающий ребра сети.

`ru.ifmo.hpc.structure.Node` — класс, описывающий узлы сети.

`ru.ifmo.hpc.structure.Graph` — класс, представляющий структуру сети, хранит информацию об узлах и ребрах между ними.

`ru.ifmo.hpc.algortithm.distribution.Distribution` — абстрактный класс, реализующий функцию подсчета средней степени вершины при данном распределении и абстрактный метод подсчета вероятности для степени вершины k .

`ru.ifmo.hpc.algortithm.distribution.DegreeDistribution` — потомок класса `Distribution`, реализующий метод вычисления вероятности для степенного распределения.

`ru.ifmo.hpc.algortithm.distribution.PoissonDistribution` — потомок класса `Distribution`, реализующий метод вычисления вероятности для распределения Пуассона.

`ru.ifmo.hpc.algortithm.generator.SequenceGenerator` — класс, генерирующий степенные последовательности на основе степенных распределений (второй шаг алгоритма).

`ru.ifmo.hpc.algortithm.generator.GraphModelGenerator` — класс, реализующий алгоритм построения сети на основе степенных последовательностей и матрицы смежности классов (третий шаг алгоритма программы).

`ru.ifmo.hpc.util.DataImporter` — класс, реализующий статические методы для импорта параметров работы алгоритма в полной (в виде степенных рядов) или сжатой (в виде параметров распределений) форме.

`ru.ifmo.hpc.util.DotLanguageExporter` — класс, предоставляющий статический метод экспорта информации о сгенерированной сети в формате языка DOT Language (четвертый шаг алгоритма).

`ru.ifmo.hpc.util.SpecialExporter` — класс, предоставляющий статический метод экспорта информации о сгенерированной сети в упрощенном формате описания графа (четвертый шаг алгоритма).

`ru.ifmo.hpc.main.ExtendedModel` — класс, являющийся точкой входа программы. Производит считывание данных и вызывает методы описанных выше классов, реализующих соответствующие шаги алгоритма.

Для обработки параметров запуска применяется свободная библиотека `Commons CLI` от `Apache Commons`.

Для запуска пакета CNM требуется в консольном приложении выполнить команду:

```
java -jar SimpleModel.jar <параметры вызова>
```

В качестве параметров программе передаются следующие значения:

- --input <имя файла> - имя файла с входными данными.
- --format <short/full> - параметр, указывающий формат входного файла:
 - short – распределение степеней вершин описывается с помощью коэффициентов степенного распределения.
 - full – распределение степеней вершин описывается с помощью степенных последовательностей.
 - alt – последовательность относительных частот степеней вершин.
- --output <имя файла> - имя файла с выходными данными.

В качестве выходных данных CNM предоставляет граф, который является результатом моделирования и описывается с помощью списков ребер графа.

3.2.2. Прикладной пакет VaccinStrategy осуществления стратегии противодействия инфекции

Прикладной пакет VaccinStrategy – программный компонент реализации стратегии противодействия распространению вируса в сети. Стратегия сводится к удалению определенного множества вершин. Из сети удаляются либо случайные вершины, либо вершины с наибольшей степенью. Пакет VaccinStrategy разработан в виде консольного приложения с использованием языка программирования Java версии 1.6. VaccinStrategy может функционировать на всех операционных системах, на которых может быть установлена платформа Java JRE 1.6.

Приложение состоит из нескольких пакетов и классов, каждый из которых выполняет специфические функции. Ниже приведены наименования классов, и описание их функций.

`ru.ifmo.hpc.structure.Node` – класс, описывающий узлы сети.

`ru.ifmo.hpc.structure.Edge` – класс, описывающий ребра сети.

`ru.ifmo.hpc.structure.Graph` – класс, представляющий структуру сети, хранит информацию об узлах и ребрах между ними в виде списков смежности.

`ru.ifmo.hpc.util.DataIO` – класс, реализующий статические методы для импорта сети и параметров работы алгоритма, а так же для экспорта результатов моделирования.

`ru.ifmo.hpc.main.Excluder` – класс, являющийся точкой входа программы. Производит считывание данных и осуществляет стратегию согласно заданным параметрам.

Для обработки параметров запуска применяется свободная библиотека Commons CLI.

Для запуска пакета VaccinStrategy требуется в консольном приложении выполнить команду:

```
java -jar VaccinStrategy.jar <параметры вызова>
```

В качестве параметров программе передаются следующие значения:

- --input <имя файла> - имя файла с входными данными.
- --config < имя файла > - имя файла с параметрами стратегии.
- --output <имя файла> - имя файла с выходными данными.

В качестве выходных данных VaccinStrategy предоставляет граф, который является результатом осуществления стратегии и описывается с помощью списков ребер графа.

3.2.3. Прикладной пакет ISM моделирования распространения вируса

Прикладной пакет ISM – программный компонент моделирования процесса распространения вируса по сети. Для моделирования процесса распространения слухов используется модель Далея-Кендалла [6].

Переменные, необходимые для описания алгоритма ДК:

I – множество вершин, которым не известна информация.

S – множество вершин, которые владеют информацией.

R – множество вершин, которые владеют информацией но не передают ее дальше.

l, a – конфигурационные параметры.

k – минимальная степень вершины, с которой начинается распространение информации.

Алгоритм моделирования процесса распространения информации ДК состоит пошагового расчета вероятности передачи информации далее. Формально эту процедуру для каждого шага можно описать следующим образом:

- $I(i) + S(j) \xrightarrow{\lambda} S(i) + S(j)$
- $S(i) + S(j) \xrightarrow{\alpha} R(i) + S(j)$
- $S(i) + R(j) \xrightarrow{\alpha} R(i) + R(j),$

где i и j – соседние вершины.

Пакет ISM разработан в виде консольного приложения с использованием языка программирования Java версии 1.6. ISM может функционировать на всех операционных системах, на которых может быть установлена платформа Java JRE 1.6.

Приложение состоит из нескольких пакетов и классов, каждый из которых выполняет специфические функции. Ниже приведены наименования классов, и описание их функций.

`ru.ifmo.hpc.structure.Node` – класс, описывающий узлы сети.

`ru.ifmo.hpc.structure.Graph` – класс, представляющий структуру сети, хранит информацию об узлах и ребрах между ними в виде списков смежности.

`ru.ifmo.hpc.util.DataIO` – класс, реализующий статические методы для импорта сети и параметров работы алгоритма, а так же для экспорта результатов моделирования.

`ru.ifmo.hpc.main.algorithm.SpreadProcess` – класс, реализующий алгоритм ДК.

`ru.ifmo.hpc.main.SpreadModel` – класс, являющийся точкой входа программы.

Производит считывание данных и вызывает методы описанных выше классов, реализующих соответствующие шаги алгоритма.

Для обработки параметров запуска применяется свободная библиотека Commons CLI.

Для запуска пакета ISM требуется в консольном приложении выполнить команду:

```
java -jar ISM.jar <параметры вызова>
```

В качестве выходных данных ISM предоставляет описание начальных условий работы алгоритма, а так же описание каждого шага распространения вируса по сети, включая детализацию по классам вершин.

3.3. Описание модулей композитного приложения в МИТП

3.3.1. Прикладной пакет CNM генерации комплексной сети

Для интеграции пакета CNM в МИТП необходимо описать его взаимодействия с платформой, опираясь на формат представления входных и выходных данных, на языке EasyPackage. Платформенный скрипт описания пакета CNM позволяет определить уровень абстракции и интерпретации в работе с входными и выходными параметрами, что обеспечивает гибкость и упрощение процедуры взаимодействия пользователя с платформой на этапе запуска задания. Платформенный скрипт описания пакета CNM определяется в соответствии с листингом 3.2.

Листинг 3.2. Скрипт описания пакета CNM

```
class DistubExtractor < TextExtractor
  def extract(str)
    distr = str.to_s
    Params = {}
    distr.strip!
```

```

        Params["Distribution"] = distr
        return Params
    end
end

class ParamsExtractor < TextExtractor
  def extract(str)
    tmpS = str.to_s
    tmpS.delete! "\r"
    lines = tmpS.split("\n")
    Params = {}
    if (lines.count == 2)
      Params["agentNum"] = lines[0].to_i

      parts = lines[1].split(" ")

      sum = 0.0
      i = 0
      parts.each { |part|
        sum += (part.to_f * i)
        i += 1
      }
      Params["PoissonKoeff"] = sum
    end
    if (lines.count > 2)
      toSkip = lines[0].to_i + 1
      Params["agentNum"] = lines[1].to_i
      lines.each { |line|
        if (toSkip == -1)
          Params["PoissonKoeff"] = line.split(" ")[1].to_f
        end
        toSkip -= 1
      }
    end

    return Params
  end
end

name "CNM"
#version nil
#display_as "Complex network modelling"
#vendor "Itmo"
#url "http://escience.ifmo.ru/"
#license "GPLv3"
#description "CNM help to understand and analyse how i.e. rumors are being
spread."
#logo ""

inputs {
  #for model time measurement
  public param {
    name "PoissonKoeff"
    type float
    #validator {|val, ctx| val > 0.0}
  }
  #for model time measurement
  public param {
    name "agentNum"
    type int
    #validator {|val, ctx| val > 0}
  }
}

```



```

public param {
  name "in_format"
  required
  display "Input Data format: short | full | alt"
  type enum ["short", "full","alt"]
  default "short"
}

public param {
  name "NodesCount"
  depends ["in_format"]
  enabled { |ctx|
    ctx.in_format == "alt"
  }
  type int
}

public param {
  name "Distribution"
  display "Distribution"

  depends ["inData?", "in_format"]
  enabled { |ctx|
    ctx.in_format == "alt"
  }
  type string
}

public file {
  name "inData"
  display "Input data file"
  extractor DistubExtractor.new
}

public file {
  name "inDataFile"
  path "/"
  depends ["NodesCount?", "Distribution?"]
  package
  filename "input.dat"
  extractor ParamsExtractor.new
  assembler erb_template rtext("cnm_alt.erb")
}

cmdline { |ctx| "java -Xms512m -Xmx1024m -jar {0} --input input.dat --
output output.dat --format " + ctx.in_format.to_s }
}

outputs {
  public file {
    name "outputFile"
    expected
    path "/"
    filename "output.dat"
  }
}

models {
  coeffs ({"Perf" => 0.0117})

  estimator { |ctx, exec, coef|
    ((ctx.PoissonKoeff * ctx.agentNum) ** 1.12) * coef.Perf / 1000
  }
}

```

```
    }  
}  
prepare_package
```

Как видно из листинга 3.2, входные параметры описаны со следующими дополнительными атрибутами:

«необходимый» (required) – данный атрибут отражает условие присутствие этого параметра при запуске, в описания пакета CNM это – `in_format`;

«значение по умолчанию» (default) – определяет начальное значение, в случае `in_format` – это `short`;

«тип» (type) – необходим для отсеивания низкоуровневых ошибок связанных с типизацией данных. В CNM параметрами с данным атрибутом являются `in_format`, `NodesCount`, `Distribution` .

«зависит» (depends) – необходим для определения данного параметра от каких-либо других. В CNM параметры, определенные с данным атрибутом, это `NodesCount`, `Distribution`, `inDataFile`.

«доступен» (enabled) – необходим для включения-выключения параметра, в зависимости от каких-либо условий. В CNM параметрами с данным атрибутом являются `NodesCount`, `Distribution`.

«пакет» (package) – используется только для файлов и показывает, что данный файл является входным для пакета. В CNM параметром с таким атрибутом является `inDataFile`.

Выходные параметры пакета описаны со следующими атрибутами:

«ожидаемый» (exrected) – используется только для файлов и показывает, что данный файл является выходным для пакета. В CNM параметром с таким атрибутом является `outputFile`.

Помимо представленных параметров на примере приведенного выше скрипта используются экстракторы (`extractor`) и ассемблер (`assembler`). Экстракторы используются для извлечения одного или нескольких параметров и подразумевает разбор входного файла WF. Ассемблер же используется для формирования файла по шаблону (внешнему файлу); он подставляет на место указанных спецификаций значения параметров.

3.3.2. Прикладной пакет *VaccinStrategy* осуществления стратегии противодействия инфекции

Для интеграции пакета *VaccinStrategy* в МИТП необходимо описать его взаимодействия с платформой, опираясь на формат представления входных и выходных данных, на языке *EasyPackage*. Платформенный скрипт описания пакета *VaccinStrategy* позволяет определить уровень абстракции и интерпретации в работе с входными и выходными параметрами, что обеспечивает гибкость и упрощение процедуры взаимодействия пользователя с платформой на этапе запуска задания. Платформенный скрипт описания пакета *VaccinStrategy* определяется в соответствии с листингом 3.3.

Листинг 3.3. Скрипт описания пакета *VaccinStrategy*

```
name "VaccinStrategy"
#version nil
display_as "Node excluding"
vendor "ITMO"
url "http://escience.ifmo.ru/"
license "GPLv3"
description "Excluder is used to exclude node(s) from network."
#logo ""

inputs {
  public file {
    name "inDataFile"
    filename "input.dat"
    path "/"
    package
    required
  }

  public file {
    name "inConfigFile"
    required
    filename "config.dat"
    path "/"
    package
  }

  cmdline { |ctx| "java -Xmx512m -jar {0} --input input.dat --config
config.dat --output output.dat" }
}

outputs {
  public file {
    name "outputFile"
    expected
    path "/"
    filename "output.dat"
  }
}

prepare_package
```

Как видно из листинга 3.3, входные параметры описаны со следующими дополнительными атрибутами:

«необходимый» (required) – данный атрибут отражает условие присутствие этого параметра при запуске, в описания пакета CNM это – inDataFile и inConfigFile;

«пакет» (package) – используется только для файлов и показывает, что данный файл является входным для пакета. В CNM параметрами с таким атрибутом являются inDataFile и inConfigFile.

Выходные параметры пакета описаны со следующими атрибутами:

«ожидаемый» (expected) – используется только для файлов и показывает, что данный файл является выходным для пакета. В CNM параметром с таким атрибутом является outputFile.

3.3.3. Прикладной пакет ISM моделирования распространения вируса

Для интеграции пакета ISM в МИТП необходимо описать его взаимодействия с платформой, опираясь на формат представления входных и выходных данных, на языке EasyPackage. Платформенный скрипт описания пакета ISM позволяет определить уровень абстракции и интерпретации в работе с входными и выходными параметрами, что обеспечивает гибкость и упрощение процедуры взаимодействия пользователя с платформой на этапе запуска задания. Платформенный скрипт описания пакета ISM определяется в соответствии с листингом 3.4.

Листинг 3.4. Скрипт описания пакета ISM

```
name "ISM"
#version nil
display_as "Information spreading modelling"
vendor "Itmo"
url "http://escience.ifmo.ru/"
license "GPLv3"
description "ISM help to understand and analise how i.e. rumors are being
spread."
#logo ""

inputs {
    public file {
        name "inDataFile"
        required
        filename "input.dat"
        path "/"
        package
    }

    public file {
        name "inConfigFile"
        required
        filename "config.dat"
```

```
        path "/"
        package
    }

    cmdline { |ctx| "java -jar {0} --input input.dat --config config.dat --
output output.dat" }
}

outputs {
    public file {
        name "outputFile"
        path "/"
        filename "output.dat"
        expected
    }
}

prepare_package
```

Как видно из листинга 3.4, входные параметры описаны со следующими дополнительными атрибутами:

«необходимый» (required) – данный атрибут отражает условие присутствие этого параметра при запуске, в описания пакета ISM это – inDataFile и inConfigFile;

«пакет» (package) – используется только для файлов и показывает, что данный файл является входным для пакета. В ISM параметрами с такими атрибутами являются inDataFile и inConfigFile.

Выходные параметры пакета описаны со следующими атрибутами:

«ожидаемый» (expected) – используется только для файлов и показывает, что данный файл является выходным для пакета. В ISM параметром с таким атрибутом является outputFile.

4. ИСПОЛЬЗУЕМЫЕ ТЕХНИЧЕСКИЕ СРЕДСТВА

ПС функционирует в рамках распределенной среды облачных вычислений под управлением многофункциональной инструментально-технологической платформы (МИТП) CLAVIRE RU.СНАБ.80066-06. Для использования ПС необходима рабочая станция с подключением к Интернет, со следующими минимальными характеристиками:

- архитектура процессора – x86, x86_64, IA64;
- объем оперативной памяти – 1 ГБ;
- объем свободного пространства на жестком диске – 1 ГБ;
- тактовая частота процессора – 1 ГГц.

Для работы с композитным приложением необходимо использовать браузеры Mozilla FireFox (версия 3.0 и выше), Google Chrome (версия 13 и выше), Opera (версия 9.0 и выше) и Internet Explorer (версия 7.0 и выше)

5. ВЫЗОВ И ЗАГРУЗКА

Исполнение ПС выполняется средствами МИТП CLAVIRE. При запуске композитное приложение описывает следующий процесс, реализуемый в среде облачных вычислений средствами МИТП:

1. По заданному перечню параметров запуска (в тексте программы на языке EasyFlow) определяются входные данные пакетов.
2. На основе мониторинга доступных вычислительных ресурсов МИТП строит оптимальное (с точки зрения минимизации общего времени выполнения) расписание исполнения цепочки запусков.
3. Экземпляр пакета генерации комплексной сети CNM запускается на исполнение на выбранном вычислительном ресурсе.
4. По окончании расчета экземпляра пакета CNM выходные данные (в форме файлов) передаются на вход пакету удаления узлов сети VaccinStrategy, несколько экземпляров которого параллельно запускаются на исполнение на подготовленных вычислительных ресурсах.
5. По окончании работы всех экземпляров пакета VaccinStrategy, выходные данные передаются такому же количеству экземпляров пакета моделирования распространения информации по сети ISM, которые запускаются на исполнение на подготовленных вычислительных ресурсах.
6. После завершения работы пакета ISM выходные файлы передаются пакету SciLab, который на основании полученных данных проводит визуализацию интегральных характеристик посредством построения графиков.
7. После окончания работы приложения выходные данные передаются в хранилище данных, и пользователь уведомляется средствами МИТП об успешном выполнении задания.

На рисунке 4.1 показан CWF в процессе исполнения композитного приложения под управлением МИТП.

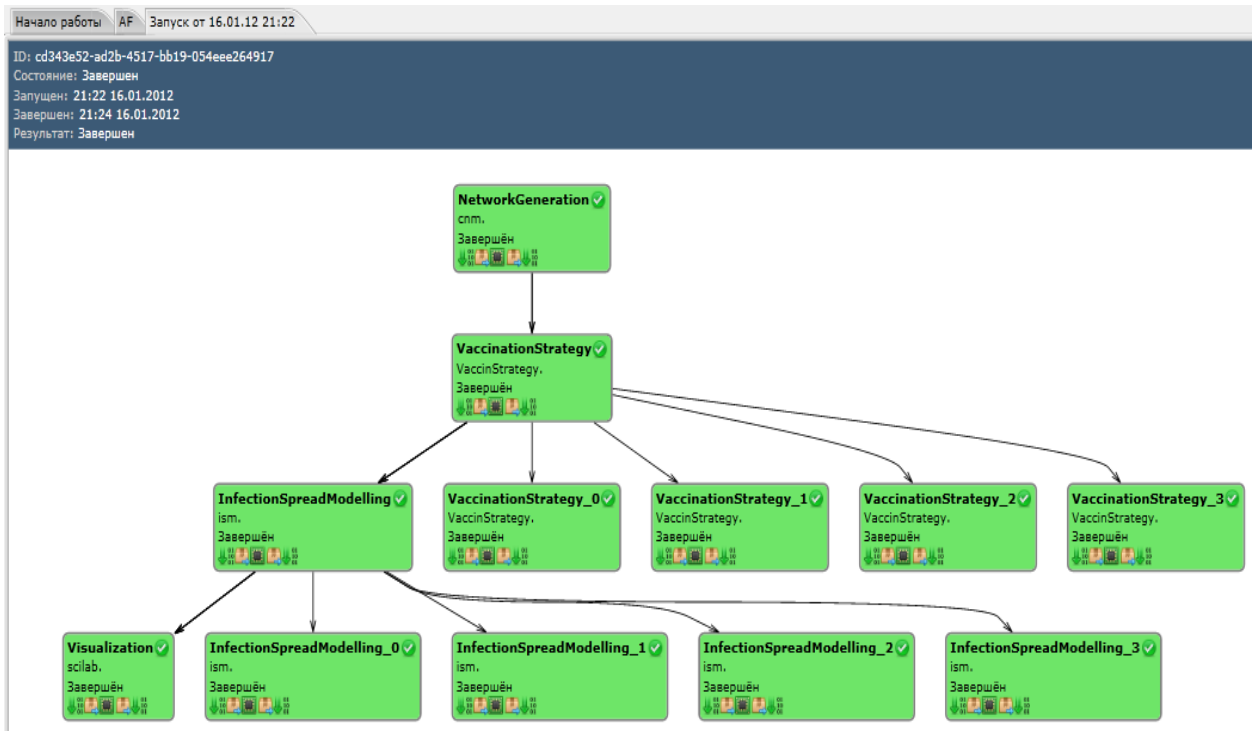


Рисунок 4.1 – Образ WF исполнения композитного приложения

6. ВХОДНЫЕ ДАННЫЕ

Входными данными для композитного приложения являются: входной файл для пакета CNM, config-файл для пакета VaccinStrategy и config-файл для пакета ISM. В качестве входного для пакета CNM использовался файл формата *short*.

- Первая строка: одно натуральное число q – число классов вершин.
- Вторая строка: одно натуральное число n – число вершин.
- q строк по q вещественных чисел разделенных пробелами – элементы матрицы E .
- q строк с двумя значениями в каждой. Первое значение – обозначение закона распределения степеней данного класса (p – Пуассона, d – степенной). Второе значение – коэффициент распределения степеней.

Config-файл для пакета VaccinStrategy имеет следующий формат:

- Первая строка: число. Если необходимо указать конкретное количество удаляемых узлов, то используется целое число (например, 105), если же необходимо удалять процент вершин, то используется дробное число (например, 0.05, указывает на то, что удаляется 5% от общего числа вершин).

- Вторая строка (необязательна): если в данной строке указано слово «*selected*», то из сети будут удаляться вершины, имеющие наибольшую степень, если же эта строка не используется, то вершины удаляются случайным образом.

Config-файл для пакета ISM имеет следующий формат: строка чисел, разделенных пробелами. Первое число: параметр λ - вероятность с которой вирус передается соседнему узлу. Второе число: параметр α - вероятность, с которой узел становится невосприимчивым к инфекции. Третье число: минимальная степень вершины, с которой начинается распространение вирусной инфекции.

7. ВЫХОДНЫЕ ДАННЫЕ

7.1. Прикладной пакет CNM генерации комплексной сети

В качестве выходных данных CNM предоставляет граф, который является результатом моделирования и описывается с помощью списков ребер графа.

Формат вывода:

- Узлы идентифицируются числами.
- Если между узлами есть ребро, то в выходной файл записывается информация вида: номер_первой_вершины--номер_второй_вершины.
- Если соединены узлы одного класса, то информация об этом ребре записывается в секцию выходного файла, начинающуюся с записи вида `cluster_i`, где i - номер класса которому эти узлы принадлежат.
- Если соединены узлы разных классов, то информация о ребре между ними записывается в секцию выходного файла, начинающуюся с записи вида `cluster_cross`.
- Если узел связан только с узлами других классов, то для того чтобы обозначить принадлежность этого узла определенному классу, в информацию о подграфе этого класса заносится индекс этой вершины.

Пример:

```
cluster_0
1--2
2--4
7
cluster_1
3--5
```


3--6

cluster_cross

7--5

1-3

7.2. Прикладной пакет *VaccinStrategy*

В качестве выходных данных *VaccinStrategy* предоставляет граф, получающийся в результате реализации стратегии противодействия. Формат описания графа полностью повторяет формат выходных данных пакета CNM.

7.3. Прикладной пакет *ISM*

В качестве выходных данных *ISM* предоставляет описание начальных условий работы алгоритма, а так же описание каждого шага распространения информации по сети, включая детализацию по классам вершин.

Формат вывода:

- Первая строка, три натуральных числа, разделенных пробелами:
 - число классов вершин исходном графе n ;
 - номер класса вершины с которой началось распространение информации;
 - степень этой вершины.
- Далее следуют строки, описывающие информацию о каждом шаге алгоритма. В каждой строке указываются следующие данные:
 - Первое число – номер шага алгоритма;
 - $(n + 1)$ группа чисел, описывающая число зараженных (R), незараженных (I) и невосприимчивых (R) вершин для всех классов (первая группа) и для каждого класса в отдельности (еще n групп).

Пример (отрывок):

```
2 0 27
1 9998 0 1 3223 0 0 6775 0
13 9986 0 1 3223 0 12 6763 0
35 9957 7 21 3202 1 14 6755 6
23 9934 42 0 3202 22 23 6732 20
46 9897 56 37 3165 22 9 6732 34
41 9856 102 2 3163 59 39 6693 43
10 9846 143 10 3153 61 0 6693 82
```

7.4. *Пакет визуализации SciLab*

На рис. 7.1 приведен пример выходного файла после работы пакета SciLab в составе данного композитного приложения.

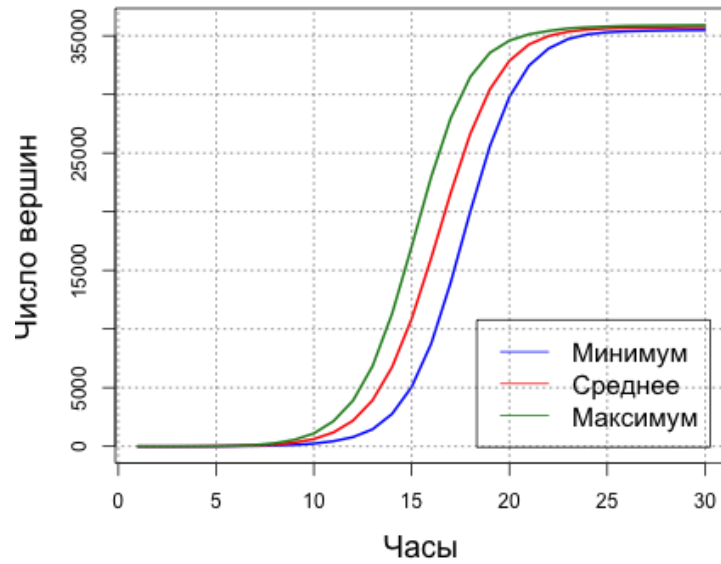


Рисунок 7.1 - Пример выходного файла пакета SciLab

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ

МИТП Многофункциональная инструментально-технологическая платформа

ПЕРЕЧЕНЬ ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- [1] Newman M.E.J. The structure and function of complex networks // *SIAM Review*, Vol. 45, N 2, pp. 167–256, 2003.
- [2] Wen-Jie Bai, Tao Zhou, Bing-Hong Wang. Interplay between HIV/AIDS Epidemics and Demographic Structures Based on Sexual Contact Networks // *arXiv:physics/0602173v1*, 2006.
- [3] Liljeros F., Edling C., Amaral L., Stanley E., Aberg Y. The Web of Human Sexual Contacts // *Nature* 411, 907-908, 2001.
- [4] Hakimi S.L. On the realizability of a set of integers as degrees of the vertices of a simple graph. // *SIAM Appl. Math.* 10, 1962, pp 496–506.
- [5] Havel V. A remark on the existence of finite graphs. (Czech), *Casopis Pest. Mat.* 80, 1955, pp. 477–480.
- [6] Daley D., Kendall D. Epidemics and rumours // *Nature*. 1964.

