ПРИМЕНЕНИЕ МЕТОДОЛОГИИ АНАЛИЗА ФОРМАЛЬНЫХ ПОНЯТИЙ ДЛЯ АВТОМАТИЧЕСКОГО ИЗВЛЕЧЕНИЯ ПРИЧИННО-СЛЕДСТВЕННЫХ СВЯЗЕЙ КОМПЬЮТЕРНЫХ АТАК

Выполнила: Буркова Надежда
ФИТ НГУ гр.0204
Кафедра общей информатики
Научный руководитель:
к.ф. – м.н., доцент НГУ
Яхъяева Г.Э

ВВЕДЕНИЕ

Компьютерные преступления в России с каждым годом совершаются все чаще, по данным МВД РФ России за 2012 год более 31,4 миллионов человек стали жертвами киберпреступников, а за 2013-й их число увеличилось на 8,6%. По оценкам ряда исследований каждую секунду в мире жертвами киберпреступников становятся 12 человек, и эта цифра с каждым годом растет. Именно поэтому необходимо предотвращать прецеденты компьютерных атак как можно раньше, а для этого необходимо анализировать большое количество информации о прецедентах и об их признаках, чтобы была возможность быстро отреагировать на проблему и принять правильные контрмеры.

ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

Цель дипломной работы - разработка модуля автоматического извлечения причинно-следственных связей из разных прецедентов компьютерных атак.

Задачи дипломной работы:

- провести анализ языков описания онтологий и выбрать наиболее подходящий для работы;
- построить онтологию предметной области компьютерной безопасности на выбранном языке;
- заполнить базу знаний прецедентами, используя данные из базы, реализованной в системе RiskPanel;
- провести анализ алгоритмов построения базиса импликаций и выбрать наиболее подходящий;
- программно реализовать модуль извлечения закономерностей между признаками прецедентов компьютерных атак.

ПРЕЦЕДЕНТЫ КОМПЬЮТЕРНЫХ АТАК

Прецедент – некоторая модель компьютерной атаки. Каждый прецедент описывается алгебраической системой $\mathfrak{A} = < \boldsymbol{A}, \boldsymbol{\sigma} >$, где $\boldsymbol{A}$ – основное множество алгебраической системы, а сигнатура $\boldsymbol{\sigma}$ – множество понятий, с помощью которых описывается предметная область компьютерной безопасности

Каждый прецедент описывается следующими понятиями:

- симптомы;
- уязвимости;
- угрозы;
- контрмеры;
- потери;
- последствия.

ОПИСАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ

Для описания предметной области была построена онтология на языке OWL, который позволяет описать прецеденты компьютерных угроз и отношения между ними. OWL является рекомендованным консорциумом Всемирной паутины. OWL пригоден для описания не только веб-страниц, но и любых объектов действительности.

В качестве редактора онтологии и фреймворка для построения базы прецедентов использовался Protégé, совместимый со стандартом W3C. Встроенный решатель Fact++ позволил проверить онтологию на непротиворечивость. В базу знаний о прецедентах компьютерных атак были автоматически занесены данные из базы, реализованной в системе RiskPanel.

АНАЛИЗ ФОРМАЛЬНЫХ ПОНЯТИЙ

Определение. ***Формальный контекст*** $K = (G, M, I)$ состоит из двух множеств G и M и отношения обладания признаком I. Элементы G называются ***объектами***, а элементы M ***признаками***.

Для произвольных подмножеств $A \subseteq G$, $B \subseteq M$ определены следующие операторы:

$A' := \{m \in M \mid gIm \; \forall g \in A\}$, $B' := \{g \in G \mid gIm \; \forall m \in B\}$.

Определение. ***Формальным понятием*** контекста $K = (G, M, I)$ называется пара вида (A, B), где $A \subseteq G$, $B \subseteq M$, $A'=B$ и $B'= A$. Множество A называется ***объемом*** понятия, а множество B его ***содержанием***.

Определение. Для контекста $K = (G, M, I)$ признаковая зависимость $A \rightarrow B$ называется ***(признаковой) импликацией*** тогда и только тогда, когда $A' \subseteq B'$ (т.е. все объекты, обладающие всеми признаками из A, обладают всеми признаками из B).

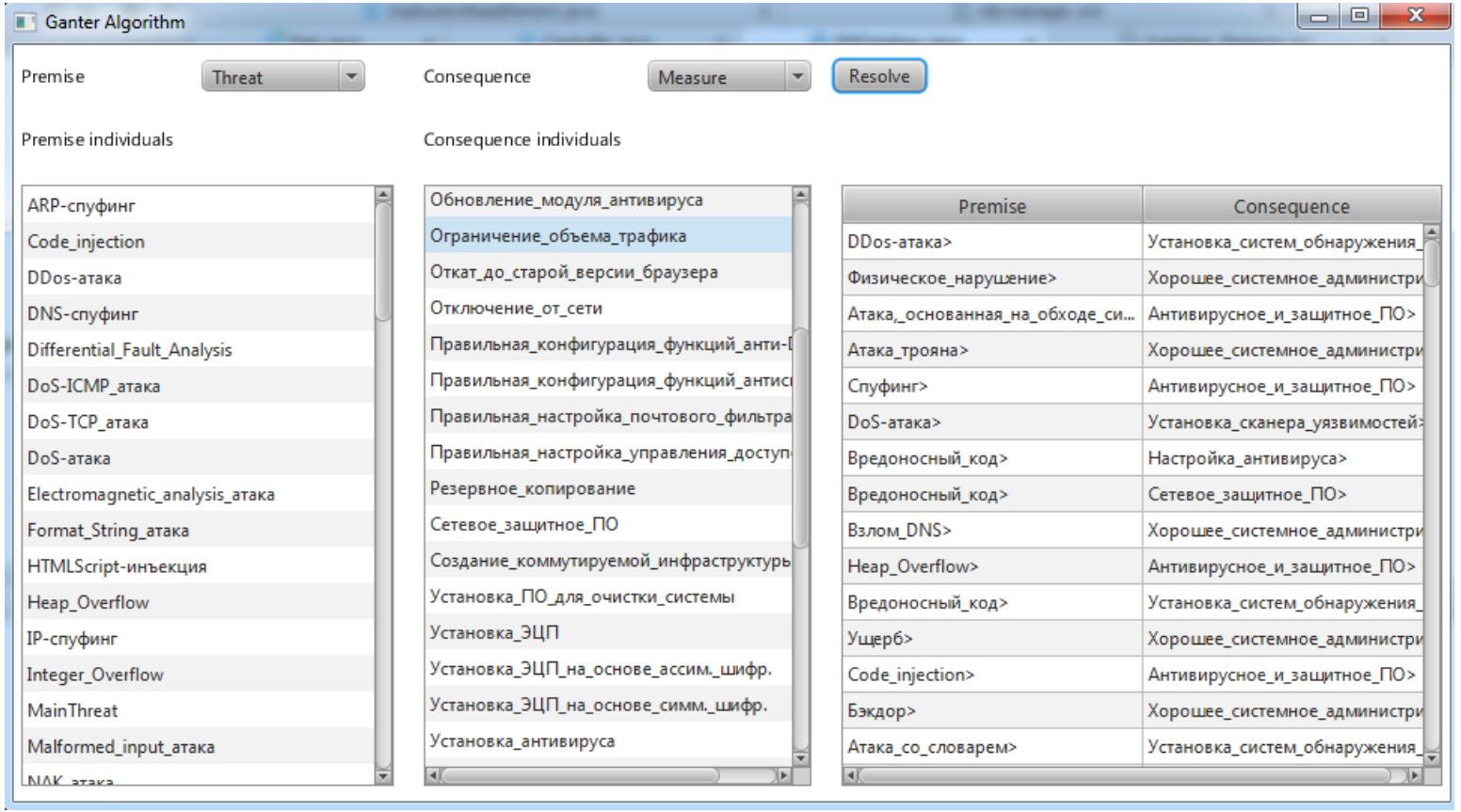
Определение. Базис импликаций – набор импликаций, из которого все остальные импликации могут быть выведены по правилам Армстронга.

АЛГОРИТМЫ ВЫЧИСЛЕНИЯ МИНИМАЛЬНОГО БАЗИСА

Существует множество различных алгоритмов для построения базиса импликаций, в данной работе были рассмотрены следующие из них:

- алгоритм Гантера (вычисляя оператор замыкания, порождает все множество псевдосодержаний вместе с множеством содержаний);
- алгоритм Объедкова-Дюкена (начиная с пустого множества признаков, добавляет новые признаки по одному и пересчитывает текущее множество псевдосодержаний вместе с множеством содержаний);
- алгоритм Бабина (приближенный базис).

РАБОТА МОДУЛЯ



РЕЗУЛЬТАТЫ

В ходе проведенной работы были получены следующие результаты:

- построена онтология на языке OWL, описывающая предметную область компьютерных атак, и подготовлена база знаний;
- изучены и применены методы анализа формальных понятий для построения минимального базиса импликаций;
- разработан модуль, автоматически извлекающий связи между признаками прецедентов компьютерных атак;
- проведено тестирование модуля на подготовленном тестовом наборе прецедентов.

Данная работа была представлена на МНСК-2014 в секции «Информационные технологии».