НАУЧНОЕ ИЗДАНИЕ МГТУ ИМ. Н. Э. БАУМАНА
НАУКА и ОБРАЗОВАНИЕ
Эл № ФС77 - 48211. Государственная регистрация №0421200025. ISSN 1994-0408
электронный научно-технический журнал

Теоретико-игровой подход к оценке рисков и нахождению уязвимостей в сетях передачи информации

 # 08, август 2013
DOI: 10.7463/0813.0630132
Басараб М. А., Вельц С. В.
УДК. 004.056

Россия, МГТУ им. Н.Э. Баумана
bmic@mail.ru
svelts@rambler.ru

Ключевыми задачами при управлении информационной безопасностью (ИБ) [1] являются идентификация угроз, уязвимостей, рисков и их оценка. Важность этих задач обусловлена тем, что результаты их решения служат входными данными для выбора контрмер по обработке рисков. Без объективной оценки рисков нельзя достичь адекватного уровня безопасности и удовлетворить требования безопасности.

В литературе [2, 3] упоминается 2 подхода к оценке рисков: качественный и количественный. В рамках сертификации CISSP[3] предлагается количественно оценивать риск, как ожидаемые потери $R = L \cdot P$, где L — оценка денежных потерь при осуществлении угрозы, P — оценка вероятности осуществления угрозы. При этом в CISSP утверждается, что количественный подход к оценке рисков — это дорогой, сложный процесс, требующий больших затрат человеческих ресурсов и времени. Например,

В данной статье предлагается подход на основе теоретико-игрового понятия ожидаемой полезности и ориентированных графов. Данный подход позволяет автоматизировать процесс оценки рисков и обнаружения наиболее уязвимых мест в системе. Также он позволяет оценивать эффект от различных мер по обработке рисков.

Применение теории игр в задачах ИБ широко освещено в литературе. Также существует масса публикаций по изучению эпидемиологических процессов в сложных сетях (напр. см. [9] и [10]). Новизна данной работы состоит в совместном практическом применении данных дисциплин на основе алгоритма информированного поиска A*. Преимуществом предлагаемого подхода по отношению к существующим (напр. моделирование ИБ на основе

байесовских сетей) является вычислительная эффективность, что позволяет применять поход к моделям с большим количеством элементов, и простота задания параметров модели экспертом.

Модель информационной системы

Рассмотрим ориентированный граф *G = (V, E)*. Множество вершин *V* представляет объекты рассматриваемой информационной системы (это могут быть компьютеры, сетевые устройства, каналы связи, люди и т. д.), которые могут подвергаться атакам. Множество рёбер *E* представляет возможные атаки/каналы утечки информации между объектами системы. Примеры таких графов приведены на рисунках 1-3.

На основе этой модели можно поставить следующие вопросы:

1. Какая атака на систему является оптимальной с точки зрения злоумышленника?
2. Каковы наиболее вероятные атаки при заданном количестве ресурсов у злоумышленника?
3. Какова вероятность успешной атаки на определённую вершину?
4. Как лучше изменить систему, чтобы снизить эту вероятность?
5. Какова вероятность компрометации узла B, при начале атаки с узла A (см. рис.3)?

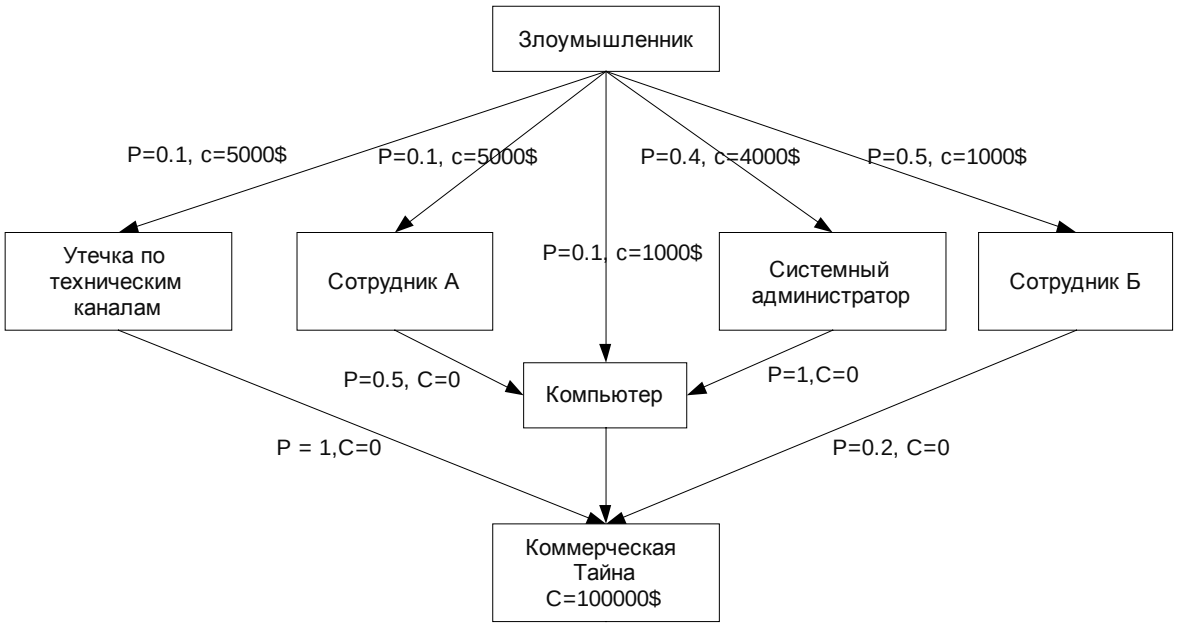

Рис. 1: Пример модели информационной системы компании. Вершины описывают объекты участвующие в информационной системе, рёбра – атаки/каналы передачи информации. Числа на рёбрах соответствуют вероятности успешной атаки и цене проведения атаки.

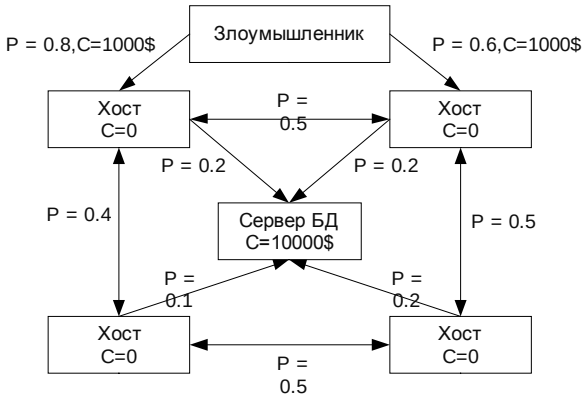

Рис. 2: Модель информационной системы на основе циклического графа.

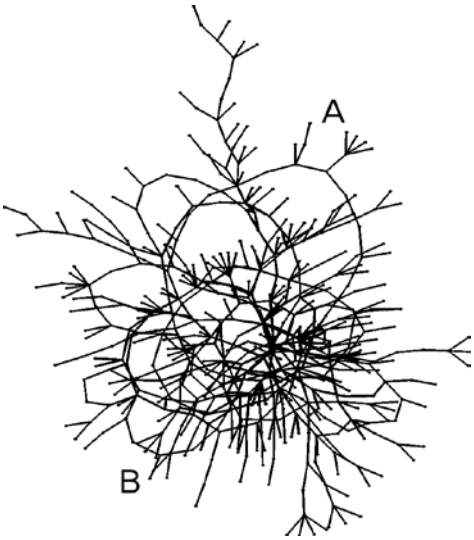

Рис. 3. Модель широкомасштабной сети на основе случайного графа.

Под оптимальностью далее понимается максимизация ожидаемой полезности, т.е. $U = M[G - C]$, где G – полученная выгода, C – цена, затраченная на проведение атаки, M — математическое ожидание указанной случайной величины.

Рассмотрим на примере (см. рис. 4), как можно посчитать это значение для пути из вершины 1 в вершину 4.

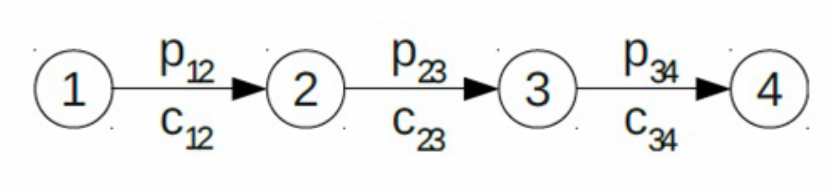

Рис. 4. Пример вычисления стоимости некоторого пути

Допустим атака начинается из вершины 1. Предпринимаем атаку на вершину 2 и платим за неё цену c_{12}, с вероятностью p_{12} мы получаем награду c_2, далее платим c_{23} за атаку и т.д. Получаем

$$u(v_1 \dots v_4) = c_1 - c_{12} + p_{12}(c_2 - c_{23} + p_{23}(c_3 - c_{34} + p_{34} c_4))$$

Или, обобщая,

$$\begin{gathered} u(v_1 \dots v_n) = c_1 - c_{12} + p_{12} u(v_2 \dots v_4) \\ u(v_i v_i) = c_i \end{gathered} \quad , \qquad (1)$$

Заметим, в общем случае может существовать несколько рёбер (т. е. несколько атак) между вершинами, петли и циклы. На вершинах определены премии c_i за компрометацию вершины.

На рёбрах определены цены c_{ij} атаки по ребру.

Следующим моментом, который необходимо обсудить для уточнения модели, является работа с рёбрами относительно времени. Возможны следующие варианты:

1. выбираем 1 атакуемое ребро на каждом шаге, т.е. осуществляем атаки последовательно;

2. выбираем K атакуемых рёбер на шаге, число K позволяет моделировать доступные нам ресурсы;

3. все рёбра обрабатываются одновременно, при этом каждой вершине сопоставлено некоторое состояние в текущий момент времени. Это наиболее общий случай.

В наиболее общей постановке задача сложна с вычислительной точки зрения, т.к. за счёт циклов может потребоваться большое количество итераций до наступления состояния равновесия. По структуре она аналогична задаче моделирования работы многослойной нейронной сети с обратными связями (либо цепи Маркова), которую можно решить с помощью алгоритма Метрополиса [4].

Далее рассмотрим частный случай, где отсутствуют циклы и выбирается 1 ребро за шаг.

Поиск оптимальной атаки

К сожалению, введённый функционал качества (1) не обладает свойством оптимальности подзадач, т.е. путь в графе, максимизирующий эту функцию, не обязательно состоит из подпутей, так же максимизирующих её. Проще всего это видеть на примере (см. рис. 5).

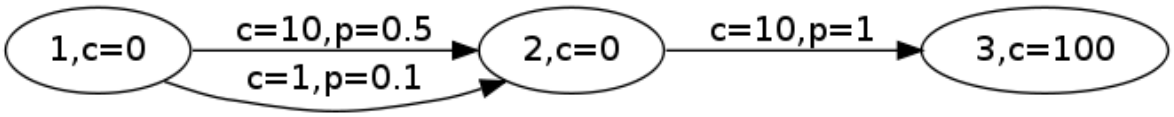

Рис. 5. Пример нарушающий свойство оптимальности подзадач

Из 1 в 3 есть 2 пути, цена первого: *-10+0,5 (-10+100)=35*, цена второго: *-1+0,1(-10+100) = 8*. Первый путь лучше, однако он содержит путь из 1 в 2, который не является локально оптимальным.

Из-за этого быстрые алгоритмы, базирующиеся на динамическом программировании (алгоритмы Дейкстры, Беллмана), оказываются неприменимы.

С другой стороны, полный перебор всех возможных путей невозможен, т.к. их может быть экспоненциальное число от количества вершин входного графа.

Если ввести подходящую эвристику, то эту проблему можно решить, применяя метод информированного поиска [6] по алгоритму SMA* (Simplified Memory-bounded A*)[6, 7].

В качестве эвристики для оценки функционала качества пути между вершинами v_1 и v_n возьмём сумму неотрицательных цен вершин и рёбер на пути.

$$\hat{u}(v_1 \dots v_n) = \max(c_n, 0) + \sum_{i=1}^{n-1} (\max(c_i, 0) + \max(-c_{i,i+1}, 0)) \quad (2)$$

Если предположить неотрицательность цен вершин (т. е. нас не штрафуют за компрометацию вершины) и рёбер (т. е. мы не получаем выгоду за сам факт атаки), то выражение (3) можно упростить до суммы цен вершин на пути:

$$\hat{u}(v_1 \dots v_n) = \sum_{i=1}^{n} c_i \quad (3)$$

Однако, сделанные такие предположения существенно ограничивают применимость модели, поэтому далее мы будем пользоваться выражением (2).

Эвристика (2) не занижает стоимость пути, следовательно является допустимой. Также (2) удовлетворяет неравенству треугольника $\hat{u}(v_1 \dots v_n) \leq \hat{u}(v_n \dots v_k) + \hat{u}(v_1 \dots v_k)$, где v_k потомок v_n, следовательно является преемственной. В соответствии с [6], поиск A* на основе такой эвристики является полным и оптимальным.

Чтобы вычислить значения эвристики во всех вершинах, можно использовать алгоритм Беллмана-Форда.

Далее обозначим $h(x) = \hat{u}(v_1, \dots, v_{n-1}, x)$ – оценка цены пути из текущего состояния x в целевое (2), $g(x)$ - цена достижения состояния x из начального (1). *f(x) = g(x) + h(x)* – оценка стоимости достижения целевого состояния по пути через вершину *x*.

Алгоритм: нахождение оптимального пути

Вход: Граф G, начальный узел start, конечный узел end

Выход: оптимальный путь

O = множество рассмотренных путей

n = количество рассмотренный путей

поместить start в O

n := 1

```
цикл:
    находим наиболее перспективный узел из уже рассмотренных
    best := argmax {x из O | f(x)}
    если best = end, то завершить вернув best
    next := следующий возможный узел в пути best
    f(next) := min(f(best), g(next) + h(next))
    если просмотрены все потомки узла best, то backup(best)
    если все потомки best в памяти, то удалить best из O
    n := n + 1
    если n > maxN то
        удалить наименее глубокий узел с мин. значением f из O
        удалить этот узел из списка последующих узлов для его
        родителя
        добавить его родителя в O, если необходимо.
        n := n – 1
    добавить next в O

function backup(n):
если n полностью обработан и имеет родительский узел, то
    f(n) := argmax {x потомок n | f(x)}
    если f(n) изменилось, то backup(parent(n))
```

Результатом предложенного подхода является наиболее привлекательный для злоумышленника путь атаки на рассматриваемую информационную систему (в предположении, что используемые нами оценки выгоды и затрат совпадают с используемыми злоумышленником). Это позволяет, во-первых, выявить наиболее слабое место в системе, во-вторых, зная путь можно перемножить вероятности реализации угроз, соответствующих рёбрам пути, и получить вероятность реализации данной атаки для её дальнейшего использования в количественных методах оценки риска.

Поиск атаки при ограниченном количестве ресурсов

В предыдущем разделе была рассмотрена задача поиска оптимальной атаки на заданный

объект в системе. Сейчас рассмотрим, как можно использовать предложенный подход для оценки вероятности атаки на систему, когда целевой объект атаки неизвестен, при условии, что ресурсы злоумышленника ограничены некоторым известным значением (предполагаем, что это значение входит в используемую модель нарушителя).

В данном случае, задачей злоумышленника является не выбор оптимального пути к целевой вершине, а выбор такого подмножества рёбер и вершин, содержащего заданную начальную вершину, что ожидаемая полезность максимальна.

Рассмотрим для каждой вершины 2 числа: c_i — математическое ожидание цены пути, ведущего от начальной вершины в данную, g_i — мат. ожидание выгоды этого пути (т.е. ожидаемая полезность, рассмотренная ранее, разделяется на 2 части).

Данная задача может быть сведена к задаче о рюкзаке (knapsack problem), если рассмотреть c_i как вес предмета, а g_i как цену предмета. Известно, что задача о рюкзаке является NP-полной, однако, если предположить, что веса целочисленные, то она допускает точное псевдо-полиномиальное решение, с помощью динамического программирования [8], которое на практике работает достаточно хорошо.

Однако это решение не будет точным для рассматриваемой задачи, т. к. после выбора одного пути цены других изменяются за счёт возможности повторно использовать выбранный путь. Так что решение, полученное в результате решения задачи о рюкзаке, переоценивает стоимость проведение атаки.

Для уточнения решения используется следующий подход:

1. Решаем задачу о рюкзаке.
2. Выбираем один из путей в качестве начального, например, максимизируя $g_i - c_i$.
3. Пересчитываем c_i и g_i для остальных путей, используя в качестве начальной вершины все вершины из выбранного в пункте 2 пути и выбирая оптимальные значения, запоминаем полученный путь.
4. Выбираем следующий путь в пункте 2 из оставшихся и повторяем пункт 3.

В результате такой жадной стратегии, мы получаем более точную оценку стоимости проведения выбранной атаки.

Выводы и направление дальнейших исследований

Была введена математическая модель на основе графов и показано, как можно на её основе

формализовать некоторые практические задачи в сфере информационной безопасности. Также был разработан эффективной способ выявления наиболее уязвимых мест в системе.

В дальнейшем полученные результаты можно развивать в следующих направлениях:

1. автоматизация оценки структуры и параметров модели на основе наблюдений, в частности, тестирования на проникновение (penetration testing) и сканирования уязвимостей;
2. разработка алгоритмов оценки системы защиты при условии параллельного проведения атак;
3. развитие методов и алгоритмов обнаружения и противодействия атакам.

Список литературы

1. BS ISO/IEC 17799:2005. Information technology — Security techniques — Code of practice for information security management.
2. ISO/IEC 27005:2008. Information technology — Security techniques — Information security risk management.
3. Bragg R. CISSP Training guide. Que Certification, 2002. 768 p.
4. Хайкин С. Нейронные сети: полный курс : пер. с англ. М.: Издательский дом «Вильямс», 2006. 1104 с.
5. Кормен Т.Х., Лейзерсон Ч.И., Ривест Р.Л., Штайн К. Алгоритмы: построение и анализ : пер. с англ. М.: Издательский дом «Вильямс», 2005. 1296 с.
6. Рассел С., Норвиг П. Искусственный интеллект: современный подход: пер. с англ. 2-е изд. М.: Издательский дом «Вильямс», 2006. 661 с.
7. Russell S. Efficient Memory-Bounded Search Methods // Proceedings of the 10th European Conference on Artificial intelligence (Vienna, Austria). New York: John Wiley & Sons, 1992. P. 1-5.
8. Martello S., Toth P. Knapsack problems. John Wesley & Sons, 1990.
9. Decision and Game Theory for Security / John S. Baras, Jonathan Katz, Eitan Altman (eds.). Springer, 2011. 259 p. (Ser. Lecture Notes in Computer Science; vol. 7037). DOI: 10.1007/978-3-642-25280-8
10. Game Theoretic Risk Analysis of Security Threats /Vicki M. Bier, M. Naceur Azaiez (eds.). Springer, 2009. 242 p. (Ser. International Series in Operations Research & Management Science; vol. 128). DOI: 10.1007/978-0-387-87767-9

SCIENTIFIC PERIODICAL OF THE BAUMAN MSTU
SCIENCE and EDUCATION
EL № FS77 - 48211. №0421200025. ISSN 1994-0408
electronic scientific and technical journal

Game-theoretic approach to risk assessment and vulnerability detection in information networks

 # 08, August 2013
DOI: 10.7463/0813.0630132
Basarab M.A., Vel'c S.V.

Bauman Moscow State Technical University, 105005, Moscow, Russian Federation
bmic@mail.ru
svelts@rambler.ru

In this paper a mathematical model based on directed graphs was presented. This model allows one to apply a quantitative approach to risk assessment and vulnerability detection in information networks, which is an important step during the security system design. A utility function and a heuristic for informed memory-bounded search SMA* algorithm were also proposed. These results could be used for security audit and countermeasures planning.

Publications with keywords: information technology security, optimization, risk assessment, computer security systems, informed search
Publications with words: information technology security, optimization, risk assessment, computer security systems, informed search

References

1. BS ISO/IEC 17799:2005. *Information technology — Security techniques — Code of practice for information security management*.

2. ISO/IEC 27005:2008. *Information technology — Security techniques — Information security risk management.*

3. Bragg R. *CISSP Training guide*. Que Certification, 2002. 768 p.

4. Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd ed. Prentice Hall, 1999. 823 p. (Russ. ed.: Haykin S. *Neironnye seti: polnyi kurs*. Moscow, Publishing House “Vil'iams”, 2006. 1104 p.).

5. Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. *Introduction to Algorithms*. 2nd ed. MIT Press and McGraw-Hill, 2001. (Russ. ed.: Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. *Algoritmy: postroenie i analiz*. Moscow, Publishing House “Vil'yams”, 2005. 1296 p.).

6. Russell S.J., Norvig P. *Artificial Intelligence: A Modern Approach*. 2nd ed. Upper Saddle River, NJ, Prentice Hall, 2003. (Russ. ed.: Rassel S.J., Norvig P. *Iskusstvennyy intellekt: sovremennyy podkhod*. Moscow, Publishing House “Vil'yams”, 2006. 661 p.) .

7. Russell S. Efficient Memory-Bounded Search Methods. *Proc. of the 10th European Conference on Artificial Intelligence*,Vienna, Austria. New York, John Wiley & Sons, 1992, pp. 1-5.

8. Martello S., Toth P. *Knapsack problems*. John Wesley & Sons, 1990.

9. Baras J.S., Katz J., Altman E., eds. *Decision and Game Theory for Security*. Springer, 2011. 259 p. (Ser. *Lecture Notes in Computer Science*; vol. 7037). DOI: 10.1007/978-3-642-25280-8

10. Bier V.M., Azaiez M.N., eds. *Game Theoretic Risk Analysis of Security Threats*. Springer, 2009. 242 p. (Ser. *International Series in Operations Research and Management Science*; vol. 128). DOI: 10.1007/978-0-387-87767-9