А. А. Малюк

ЭНТРОПИЙНЫЙ ПОДХОД К МОДЕЛИРОВАНИЮ СИСТЕМ И ПРОЦЕССОВ ЗАЩИТЫ ИНФОРМАЦИИ

Моделирование систем и процессов защиты информации является важнейшим методом реализации упреждающей стратегии защиты. Последняя основана на прогнозировании показателей уязвимости информации в различных ситуациях защиты и условиях функционирования защищаемых систем.

Разработанные на сегодня подходы в принципе позволяют решать большинство задач моделирования указанных процессов. Однако чтобы воспользоваться предлагаемыми в этой области моделями, должны быть известны функциональные зависимости значений показателей защищенности и уязвимости от большого числа параметров системы защиты, а также зависимости самих параметров от размеров ресурсов, вкладываемых в реализацию систем и процессов защиты.

Так как на практике из-за отсутствия необходимых статистических данных не удается строго выполнить эти условия, то предлагаемые формальные модели могут применяться только в совокупности с неформальными методами анализа и прогнозирования. Таким образом, моделирование представляется в виде алгоритма автоформализации знаний эксперта-аналитика [1].

Если изначально опираться на технологию автоформализации знаний, то для увязки показателей уязвимости с ресурсами, используемыми для защиты, продуктивным может оказаться подход к моделированию, когда для оценки функции полезности системы обеспечения безопасности информации применяется понятие энтропии системы.

Сформулируем описание системы обеспечения безопасности информации как системы с максимальной полезностью, под которой понимается наиболее полное использование ресурсов для целей защиты информации.

Пусть $x_1, x_2, \ldots, x_n$ — некоторые показатели, характеризующие деятельность отдельных подсистем системы обеспечения безопасности информации, достижение которых сопряжено с удельными затратами $r_1, r_2, \ldots, r_n$. При этом суммарные затраты системы ограничиваются величиной соответствующего бюджета I. Для оценки максимально возможного в этом случае уровня обеспечения безопасности информации необходимо максимизировать функцию полезности

$$u = u(x_1, x_2, \ldots, x_n, I), \tag{1}$$

соответствующую принятой концепции и структуре системы безопасности информации, при бюджетном ограничении

$$\sum_{i=1}^{n} x_i r_i = I. \tag{2}$$

Определим лагранжиан L

$$L = u(x_1, x_2, \ldots, x_n, I) + \lambda (I - \sum_i r_i x_i), \tag{3}$$

где λ — множитель Лагранжа, связанный с уравнением (2).

Проводя теперь обычным способом максимизацию, получаем, что значения параметров системы обеспечения безопасности информации определяются из решения системы уравнений

$$\frac{\partial L}{\partial x_i} = \lambda r_i. \tag{4}$$

Это решение может быть записано в виде

$$x_i = x_i(r_1, r_2, \ldots, r_n, I). \tag{5}$$

Можно показать также, что при заданном уровне полезности

$$\left.\frac{\partial I}{\partial r_i}\right|_{u=\bar{u}} = x_i. \tag{6}$$

Таким образом, уравнения (1)–(6) описывают систему обеспечения безопасности информации как систему с максимальной полезностью.

Покажем, что задача максимизации функции полезности такой системы может быть сведена к максимизации ее энтропии.

А. Дж. Вильсон применил максимизацию энтропии при решении проблемы оптимизации транспортных потоков в городских системах [2]. Он показал, что энтропия в этом случае связана с распределением вероятностей:

$$S = -\sum_i \sum_j p_{ij} \ln p_{ij}, \tag{7}$$

где $\rho_{ij} = T_{ij} / T$ может интерпретироваться как распределение вероятностей (T_{ij} — число поездок из зоны i в зону j, T — полное число поездок, которое является фиксированной величиной).

Рассмотрим возможность применения данного выражения для определения энтропии системы обеспечения безопасности информации как системы с максимальной полезностью. В этом случае необходимо решить проблему соизмерения значений частных функций полезности отдельных ее подсистем. Чтобы аналог выражения (7) интерпретировался как энтропия системы обеспечения безопасности информации, должны быть введены некоторые относительные единицы, связывающие количественные характеристики деятельности отдельных подсистем с подходящей фиксированной величиной. В качестве последней может быть использована величина бюджета I. Тогда такая относительная единица будет представлена в виде

$$y_i = \frac{x_i r_i}{I}. \tag{8}$$

Теперь система с максимальной полезностью может быть описана в терминах y_i следующим образом:

$$u = u\left(\frac{y_1 I}{r_1}, \frac{y_2 I}{r_2}, \ldots, \frac{y_n I}{r_n}, I\right) \to max, \tag{9}$$

$$\sum_{i=1}^{n} y_i = 1. \tag{10}$$

Функция Лагранжа имеет вид

$$L = u + \lambda \left(1 - \sum_i y_i\right). \tag{11}$$

Дифференцирование ее по y_i приводит к следующей системе уравнений:

$$\frac{du}{dy_i} = \lambda, \; i = 1, 2, \ldots, n, \tag{12}$$

откуда

$$y_i = y_i(r_1, r_2, \ldots, r_n, I) \tag{13}$$

и

$$\left.\frac{dI}{dr_i}\right|_{u=\bar{u}} = \frac{y_i I}{r_i}. \tag{14}$$

Легко проверяется, что уравнения (9)–(14) описывают ту же систему, что и уравнения (1)–(6).

Предположим теперь, что для анализа этой системы используется принцип максимизации энтропии

$$S = -\sum_{i=1}^{n} y_i \ln y_i \tag{15}$$

при известном ограничении (10).

Ограничения по другим видам ресурсов сформулируем следующим образом:

$$f_i(y_1, y_2, \ldots, y_n) = g_j, \; j = 1, 2, \ldots, k, \tag{16}$$

где для удобства все члены, содержащие y_i, входят в f_i, а все остальные (константы) в g_i.

Функция Лагранжа в этом случае имеет вид

$$L = S + \lambda\ (1 - \sum_{i=1}^{n} y_i) + \sum_{j=1}^{k} \mu_j\ (g_j - f_j), \quad (17)$$

где λ и μ — множители Лагранжа, связанные соответственно с уравнениями (10) и (16).

Дифференцируя выражение (17) по y_i, получим систему уравнений, решая которую совместно с уравнениями (10) и (16) в итоге имеем

$$\ln y_i = -\lambda - \sum_{j=1}^{k} \mu_j \frac{df_i}{dy_i}. \quad (18)$$

Из анализа выражения (18) следует, что задача максимизации энтропии формально эквивалентна максимизации функции полезности, записанной в виде:

$$u = S + \sum_{j=1}^{k} \mu_j\ (g_j - f_j), \quad (19)$$

при ограничении (10) и очевидном условии, что параметры $\mu_j < 1$, а число ограничений k меньше числа переменных n.

Таким образом, и при максимизации энтропии, и при анализе системы с максимальной полезностью, в конце концов, будет получен один и тот же результат. Однако максимизация энтропии имеет принципиально важные для решения специфической задачи исследования состояний безопасности информации преимущества перед статистическим подходом, так как позволяет учитывать априорную информацию об отдельных ограничениях, накладываемых на y_i, а также делает возможной индивидуальную интерпретацию ограничений. Кроме того, этот подход оказывается полезным при построении динамических моделей.

Сформулируем теперь на основе рассмотренного энтропийного подхода соответствующую модель системы обеспечения безопасности информации. Пусть состояние x_i системы соответствует некоторому ресурсу $f(x_i)$. Под состоянием x_i будем понимать некоторый i-й набор средств защиты информации. При этом справедливо ограничение

$$\sum_i \rho_i f(x_i) = E[f(x)] \le U, \quad (20)$$

где ρ_i — вероятность состояния x_i;

U — лимит на ресурс либо ограничение на полезный эффект.

Тогда задача поиска оптимального с точки зрения максимизации уровня обеспечения безопасности информации распределения величины x_i формально записывается в виде:

$$S \to max, \quad (21)$$

$$\sum_i \rho_i f(x_i) = U, \quad (22)$$

$$\sum_i \rho_i = 1, \quad (23)$$

где $S = -\sum_i \rho_i \ln \rho_i$ — энтропия системы.

Решение данной задачи методом неопределенных множителей Лагранжа имеет вид:

$$\rho_i = exp\ [-\lambda - \mu f(x_i)], \quad (24)$$

где λ и μ — множители Лагранжа.

С учетом условия (23) получаем

$$e^{\lambda} = \sum_i exp\ [-\mu f(x_i)] \quad (25)$$

При этом искомое оптимальное распределение представляется в виде распределения Больцмана:

$$\rho = \frac{exp\ [-\mu f(x_i)]}{\sum_i exp\ [-\mu f(x_i)]} \quad (26)$$

Таким образом, макросостояние системы обеспечения безопасности информации задается функцией $f(x_i)$, имеющей в нашем случае смысл ресурса, и некоторым параметром μ, аналогом

температуры ($T = 1/\mu$) в физических системах. Опираясь далее на физическую аналогию (второе начало термодинамики), введем так называемую статистическую сумму

$$Z = \sum_i exp\ [-\mu f(x_i)] \tag{27}$$

а также величину $F = -(1/\mu)lnZ$ — аналог свободной энергии в физических системах. Используя их, можем получить на основе известного соотношения Г. Гельмгольца для свободной и связанной энергии следующие соотношения:

$$F = U - \frac{1}{\mu} S\ , \tag{28}$$

$$S = -\frac{dF}{d(1/\mu)}\ , \tag{29}$$

$$U = F - \frac{1}{\mu}\frac{dF}{d(1/\mu)} = \frac{d}{d\mu}(\mu F) = -\frac{d(lnZ)}{d\mu}\ . \tag{30}$$

К этим соотношениям добавляется условие монотонности возрастания U и S при положительном μ и $f(x_i) \neq const$.

Таким образом, макросостояние системы обеспечения безопасности информации можно задать четырьмя взаимосвязанными характеристиками U, F, μ и S. Их интерпретация зависит от постановки решаемой задачи, а также от особенностей конкретной исследуемой системы. В частности, F может интерпретироваться как суммарные прямые издержки системы на создание определенного уровня обеспечения безопасности, а $\frac{1}{\mu} S$ — как косвенные затраты на поддержание этого уровня.

Обобщение задачи поиска оптимального распределения на случай задания более одного вида ограничений на ресурсы формально записывается в виде:

$$S = -\sum_i \rho_i ln\rho_i \to max, \tag{31}$$

$$\sum_i \rho_i f_r(x_i) = E[f_r(x)],\ r=1, 2, ..., n, \tag{32}$$

$$\sum_i \rho_i = 1. \tag{33}$$

Аналогично (27) строится функция

$$Z(\mu_1, \mu_2, ..., \mu_m) = \sum_i exp\ [-\sum_{r=1}^{m} \mu_r f_r(x_i)]. \tag{34}$$

Тогда

$$\rho_i = exp\ [-\sum_{r=1}^{m} [\lambda + \mu_r f_r(x_i)], \tag{35}$$

где $\lambda = lnZ$.

Остальные множители Лагранжа определяются из ограничений (32) и (33), записываемых в виде:

$$E[f_r(x)] = -\frac{d}{d\mu_r} lnZ\ . \tag{36}$$

Можно вычислить максимальное значение энтропии:

$$S_{max} = \lambda + \sum_{r=1}^{m} \mu_r E[f_r(x)] \tag{37}$$

и возможные флуктуации, рассчитывая дисперсию распределения

$$\Delta^2 f_r(x) = E\ [f_r(x)]^2 - \{E[f_r(x)]\}^2 = \frac{d^2}{d\mu_r^{\,2}} lnZ. \tag{38}$$

Если задана зависимость f_r не только от x, но и от независимых параметров $\alpha_j (j = 1, 2, ..., L)$, то можно оценить значение ее производных по максимуму энтропии

$$E(\frac{df_r}{d\alpha_j}) = \frac{1}{\mu}\frac{d}{d\alpha} lnZ. \tag{39}$$

Предположим, что функции ограничений $f_r(x)$ можно менять независимым образом для всех r и i. Допустим также независимое изменение средних значений f_r. Тогда

$$\delta\lambda = \delta lnZ = -\sum_{r=1}^{m} \{\delta\mu_r E[f_r(x)] + \mu_r E[\delta f_r(x)]\}, \tag{40}$$

и, воспользовавшись (37), получим

$$\delta S = \sum_{r=1}^{m} \mu_r \{\delta E[f_r(x)] - E[\delta f_r(x)]\} = \sum_{r=1}^{m} \mu_r \delta G_r , \quad (41)$$

где параметр G_r определяется соответствующим видом ограничений и является r-м видом «теплоты», если пользоваться терминологией термодинамики, соотношения которой мы и положили в основу всех наших рассуждений. Таким образом, μ_r — весовой коэффициент при G_r, является, следовательно, r-м видом «температуры».

Из уравнения (41) легко может быть получен его частный случай

$$dS = \mu dE[f(x)] + \mu \sum_k \bar{x}_k dx_k, \quad (42)$$

где $\bar{x}$ — среднее значение обобщенной силы, действующей на внешнюю координату x_k.

Это выражение является аналогом второго закона термодинамики и описывает процесс релаксации системы обеспечения безопасности информации в равновесное состояние, определяющее ее потенциальные возможности.

Для практического применения предложенных энтропийных методов моделирования необходимо увязать макропараметры системы U, F, μ, S с конкретными характеристиками отдельных ее подсистем и элементов, что может выполнить эксперт-аналитик.

СПИСОК ЛИТЕРАТУРЫ:

1. *Громов Г. Р.* Национальные информационные ресурсы: проблемы промышленной эксплуатации. М.: Наука, 1985.
2. *Вильсон А.* Энтропийные методы моделирования сложных систем / Пер. с англ. М.: Наука, 1978.