ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ КОМПЬЮТЕРНЫХ НАУК

МАТЕМАТИЧЕСКИЕ СТРУКТУРЫ и МОДЕЛИРОВАНИЕ

Выпуск 19

Омск 2009

Журнал «**Математические структуры и моделирование**». – Омск, 2009. – Вып. 19. – 122 с.

ISBN

Адрес научной редакции

Россия, 644053, Омск - 53, ул. Грозненская, 11
Омский государственный университет
факультет компьютерных наук

E-mail: `guts@omsu.ru`
`msm@cmm.univer.omsk.su`

ISBN

МАТЕМАТИЧЕСКИЕ СТРУКТУРЫ и МОДЕЛИРОВАНИЕ

В журнале публикуются статьи, в которых излагаются результаты исследований по фундаментальной и прикладной математике, теоретической физике и размышления, касающиеся окружающей нас природы и общества.

Публикуются также статьи по информационным технологиям, компьютерным наукам, защите информации, философии и истории математики.

Объекты исследования должны быть представлены в форме некоторых математических структур и моделей.

Журнал является реферируемым. Рефераты статей публикуются в «Реферативном журнале» и в журналах «Zentralblat für Mathematik» (Германия) и «Mathematical Reviews» (США).

Электронная версия журнала представлена в сети «Интернет» по адресу:

`http://cmm.univer.omsk.su`

Журнал издается на коммерческие средства факультета компьютерных наук Омского государственного университета.

Наш E-mail:

`msm@cmm.univer.omsk.su`

Подробную информацию можно найти на Web-сервере:

`http://cmm.univer.omsk.su`

СОДЕРЖАНИЕ

Фундаментальная математика

Прикладная математика и моделирование

Теоретическая физика

Компьютерные науки и защита информации

продолжение на след.странице

**ФАКУЛЬТЕТ
КОМПЬЮТЕРНЫХ НАУК**

Омкого государственного университета

Факультет создан 1 мая 2003 г.

Первый набор студентов был осуществлен в 2001 году.

Кафедры и лаборатории. На факультете открыты:

– кафедра информационной безопасности;
– кафедра кибернетики;
– кафедра вычислительных систем;
– научно-исследовательская лаборатория программного обеспечения и компьютерных сетей;
– научно-исследовательская лаборатория социокибернетики.
– учебно-научный центр переподготовки и повышения квалификации в области защиты информации и информационной безопасности;
– сетевая академия CISCO;
– академия SUN Microsystems;
– академия Oracle.

Обучение по специальностям:

1) 075200 – Компьютерная безопасность / анализ безопасности компьютерных систем. Квалификация «математик». Срок обучения 5 лет 6 месяцев. Выпуск в феврале.

2) 220100 – Вычислительные машины, комплексы, системы и сети. Квалификация «инженер». Срок обучения 5 лет. Имеются очно-заочное отделение, а также сокращенная форма обучения (очно – 3 года, заочно – 3,5 года).

Телефон (факс): (+7 3812) 64-83-11

Математические структуры и моделирование
2009, вып. 19, с. 5–18

УДК 519.214.5

АСИМПТОТИЧЕСКОЕ РАЗЛОЖЕНИЕ В ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЕ ДЛЯ СУММ ЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

А.Г. Гринь

It is received an asymptotic expansion with a member of an order n^{-1} in the central limit theorem for stationary sequences with uniform strong mixing.

1. Введение. Формулировки основных результатов

Пусть $\{\xi_n\} = \{\xi_n,\ n = 1, 2, ...\}$ – стационарная в узком смысле последовательность и пусть

$$\mathbf{E}\xi_1 = 0,\ \mathbf{E}\xi_1^2 < \infty,\ T_n = \sum_{j=1}^{n} \xi_j,\ \sigma_n^2 = \mathbf{D}T_n \to \infty,\ n \to \infty,$$

$$F_n(x) = \mathbf{P}\{T_n < x\sigma_n\}, \quad \Phi(x) = (2\pi)^{-\frac{1}{2}} \int\limits_{-\infty}^{x} \exp\left\{-\frac{t^2}{2}\right\} dt,$$

$$\Delta_n = \sup_x |F_n(x) - \Phi(x)|\,.$$

Для последовательностей слабо зависимых величин оценки скорости стремления Δ_n к нулю последовательно улучшались на протяжении более, чем тридцати лет, пока, наконец, Е.Рио в [1] не получил неулучшаемую по порядку оценку $\Delta_n = O(n^{-\frac{1}{2}})$ для последовательностей ограниченных величин с равномерно сильным перемешиванием (φ–перемешиванием).

Настоящая работа посвящена дальнейшему уточнению центральной предельной теоремы для сумм слабо зависимых случайных величин. Выделен класс последовательностей $\{\xi_n\}$ (так называемые последовательности с симметричным распределением), для которых при экспоненциально быстром φ–перемешивании, $\mathbf{E}|\xi_1|^5 < \infty$ и

$$\limsup_{t\to\infty} |\mathbf{E}\exp\{it\xi_1\}| < 1, \tag{1}$$

доказан аналог асимптотического разложения Эссеена в центральной предельной теореме (см., например [2, гл.5, теорема 20]). Из этого разложения (см. теорему 1) в частности, следует, что $\Delta_n = O(n^{-1})$.

Омский государственный университет.
E-mail: `griniran@gmail.com`
Работа поддержана грантом РФФИ 06–01–00127.

1. Последовательности с симметричным распределением

Пусть $\xi = (\xi_1, ..., \xi_n)$ – случайный вектор, δ – отображение множества $\{1, 2, ..., n\}$ в множество $\{-1, 1\}$, то есть $\delta = (\delta_1, ..., \delta_n), \delta_k = \pm 1,\ k = 1, 2, ..., n$ и пусть $\Delta = \{\delta\} = \{-1, 1\}^{\{1,2,...,n\}}$. Обозначим $\delta * \xi = (\delta_1\xi_1, ..., \delta_n\xi_n)$. Распределение вектора ξ будем называть симметричным, если все векторы $\delta * \xi,\ \delta \in \Delta$ имеют одинаковое распределение. Будем говорить, что последовательность $\{\xi_n\}$ имеет симметричное распределение, если все ее конечномерные распределения симметричны.

Простейшим примером последовательности с симметричным распределением является последовательность независимых одинаково распределенных величин с симметричными распределениями.

Введем характеристическую функцию вектора ξ :

$$f_{\xi_1,...,\xi_n}(t_1, ..., t_n) = \mathbf{E} \exp \left\{ i \sum_{k=1}^{n} t_k \xi_k \right\}$$

Будем использовать обозначение $\xi \overset{d}{=} \eta$ в случае, когда распределения случайных векторов ξ и η совпадают, и $\{\xi_n\} \overset{d}{=} \{\eta_n\}$, когда совпадают конечномерные распределения последовательностей $\{\xi_n\}$ и $\{\eta_n\}$.

Лемма 1. *Следующие условия эквивалентны:*

a) Последовательность $\{\xi_n\}$ имеет симметричное распределение;

b) При любом натуральном n и любых действительных $t_1, ..., t_n$

$$f_{\xi_1,...,\xi_n}(t_1, ..., t_n) = \mathbf{E} \prod_{k=1}^{n} \cos(t_k \xi_k); \tag{2}$$

c) $\{\xi_n\} \overset{d}{=} \{\varepsilon_n \xi_n\}$, где $\{\varepsilon_n\}$ – последовательность независимых случайных величин таких, что

$$\mathbf{P}\{\varepsilon_n = 1\} = \mathbf{P}\{\varepsilon_n = -1\} = \frac{1}{2}$$

и последовательность случайных величин $\{\varepsilon_n\}$ не зависит от $\{\xi_n\}$.

Доказательство достаточно прозрачно и здесь не приводится.

З а м е ч а н и е 1. Если $\{\xi_n\}$ – стационарная последователность, удовлетворяющая условию РСП с коэффициентом $\varphi(n)$, а $\{\varepsilon_n\}$ – последовательность, определенная в пункте c) леммы 1, то последовательность $\{\varepsilon_n \xi_n\}$ также удовлетворяет условию РСП с коэффициентом $\varphi_1(n) \leq \varphi(n)$ [3].

З а м е ч а н и е 2. Нетрудно убедиться, что если $\{\xi_n\}$ – последовательность с симметричным распределением,то при каждом натуральном p последовательность $\left\{ \sum\limits_{l=1}^{p} \xi_{(j-1)p+l}, j = 1, 2, ... \right\}$ также имеет симметричное распределение.

Лемма 2. *Если* $\{\xi_n\}$ *–стационарная последовательность с симметричным распределением и* $\mathbf{E}|\xi_1|^4 < \infty$, *то*

a) $\mathbf{E}T_n = 0,\ \mathbf{E}T_n^3 = 0,\ n = 1, 2, ...;$
b) $\sigma_n^2 = \mathbf{D}T_n^2 = n\sigma_1^2,\ n = 1, 2, ...;$
c) $\mathbf{E}T_n^4 = n\mathbf{E}\xi_n^4 + 6 \sum\limits_{1 \le i < j \le n} \mathbf{E}\xi_i^2\xi_j^2,\ n = 1, 2, ...$

Утверждения леммы являются следствием симметричности распределений величин $\xi_j, \xi_j\xi_k), \xi_j, \xi_k^3,\ j \ne k$.

2. Вспомогательные результаты

Пусть $\{\xi_n\} = \{\xi_n,\ n = 1, 2, ...\}$ – стационарная последовательность. Обозначим через $\mathcal{F}_{\le n}$ и $\mathcal{F}_{\ge n}$ σ–алгебры, порожденные, соответственно, семействами $\{\xi_k : k \le n\}$ и $\{\xi_k : k \ge n\}$.

Лемма 3. *Пусть последовательность* $\{\xi_n\}$ *удовлетворяет условию равномерно сильного перемешивания (*φ*–перемешивания) с коэффициентом перемешивания* $\varphi(n)$ *и пусть случайные величины* ξ *и* η *измеримы относительно* $\mathcal{F}_{\le 0}$ *и* $\mathcal{F}_{\ge n}$ *соответственно,*

$$\|\xi\|_p = (\mathbf{E}|\xi|^p)^{\frac{1}{p}} < \infty, \quad \|\eta\|_q < \infty, \quad p > 1,\ q > 1, \quad p^{-1} + q^{-1} = 1.$$

Тогда при любых комплексных a *и* b

$$|\mathbf{E}\xi\eta - \mathbf{E}\xi\mathbf{E}\eta| \le 2\varphi^{\frac{1}{p}}(n)\|\xi - a\|_p\|\eta - b\|_q. \tag{3}$$

Если же

$$\|\xi\|_1 = \mathbf{E}|\xi| < \infty,\ \|\eta\|_\infty = \mathrm{vrai\,sup}|\eta| \le \infty,$$

то

$$|\mathbf{E}\xi\eta - \mathbf{E}\xi\mathbf{E}\eta| \le 2\varphi(n)\|\xi - a\|_1\|\eta - b\|_\infty. \tag{4}$$

Доказательство леммы легко получается, например, из теоремы 17.2.3 в [4] или из [5, с.236].

Лемма 4. *Пусть последовательность* $\{\xi_n\}$ *удовлетворяет* φ*–перемешивания,* $\sigma_n \to \infty,\quad n \to \infty$, *и пусть* $\mathbf{E}|\xi|^p < \infty,\ p > 2$. *Тогда* $\|T_n\|_p \le C(p)\,\sigma_n$, *где* $0 < C(p) < \infty$ *не зависит от* n.

При $2 < p < 3$ утверждение леммы доказано в [4, теорема 18.5.1], в общем случае оно следует, например, из теоремы 1.1 в [6].

Обозначим $\gamma_k\,(\xi)$ $k-$й семиинвариант случайной величины ξ.

Лемма 5. *Пусть* $\{\xi_n\}$ *–стационарная последовательность с симметричным распределением пусть* $\mathbf{E}|\xi_1|^4 < \infty$. *Тогда*

a) $\gamma_1(T_n) = 0,\ \gamma_3(T_n) = 0,\ n = 1, 2, ...;$
b) $\gamma_2(T_n) = n\sigma_1^2,\quad n = 1, 2, ...;$

c) Если $\sum\limits_{k=1}^{\infty} \varphi^{\frac{1}{2}}(k) < \infty$, *то*

$$\gamma_4 = \gamma_4(\xi_1) + 6 \sum_{k=1}^{\infty} \mathbf{E} \left(\xi_0^2 - \mathbf{E}\xi_0^2\right) \left(\xi_k^2 - \mathbf{E}\xi_k^2\right) < \infty.$$

Если же $\sum\limits_{k=1}^{\infty} k\varphi^{\frac{1}{2}}(k) < \infty$, *то*

$$\sup_n |\gamma_4(T_n) - n\gamma_4| \leq C < \infty.$$

Д о к а з а т е л ь с т в о.

Так как в наших предположениях $\gamma_k(T_n) = \mathbf{E}T_n^k,\ k = 1, 2, 3$, то утверждения a) и b) следуют из соответствующих утверждений леммы (2). Обозначим $\rho_k = \mathbf{E}\left(\xi_0^2 - \mathbf{E}\xi_0^2\right)\left(\xi_k^2 - \mathbf{E}\xi_k^2\right)$. В силу леммы 3 $|\rho_k| \leq 2 \left\|\xi_0^2 - \mathbf{E}\xi_0^2\right\|_2^2 \varphi^{\frac{1}{2}}(k)$, откуда следует первое утверждение пункта c) леммы. С учетом утверждений b) и c) леммы (2) получаем

$$\gamma_4(T_n) = \mathbf{E}T_n^4 - 3\left(\mathbf{E}T_n^2\right)^2 = n\mathbf{E}\xi_n^4 + 6 \sum_{1 \leq i < j \leq n} \mathbf{E}\xi_i^2\xi_j^2 - 3n^2\sigma_1^4 =$$

$$= n\mathbf{E}\xi_1^4 - 3n\sigma_1^4 + 6 \sum_{1 \leq i < j \leq n} \mathbf{E}\left(\xi_i^2 - \mathbf{E}\xi_i^2\right)\left(\xi_j^2 - \mathbf{E}\xi_j^2\right) =$$

$$= n\gamma_4(\xi_1) + 6 \sum_{k=1}^{n-1} (n-k)\rho_k.$$

Отсюда

$$|\gamma_4(T_n) - n\gamma_4| \leq 6n \sum_{k=n}^{\infty} |\rho_k| + 6 \sum_{k=1}^{n-1} k|\rho_k| \leq$$

$$\leq 12 \left\|\xi_0^2 - \mathbf{E}\xi_0^2\right\|_2^2 \sum_{k=1}^{\infty} k\varphi^{\frac{1}{2}}(k) < \infty.$$

Лемма доказана.

Пусть $\mathbf{E}\xi_1^2 < \infty$, а n и p–натуральные числа. Обозначим

$$Q_j = {\sigma_n}^{-1} \sum_{l=1}^{p} \xi_{(j-1)p+l}, \quad j = 1, 2, ... \tag{5}$$

$$f_r(t) = \mathbf{E}\exp\left\{it \sum_{j=1}^{r} Q_j\right\},\ r = 1, 2, ...$$

Введем функцию

$$F_k(z) = \mathbf{E} \prod_{j=1}^{k} \left(f_1 + z\left(\exp\left\{itQ_j\right\} - f_1\right)\right) =$$

$$= \sum_{m=1}^{k} z^m f_1^{k-m} \mathbf{E} S_m(X_1, ..., X_k),$$

где $f_1 = f_1(t),\ X_j = \exp\{itQ_j\} - f_1,\ j = 1, ..., k$, а

$$S_0(x_1, ..., x_k) = 1, \quad S_m(x_1, ..., x_k) = \sum_{1 \le j_1 < ... < j_m \le k} x_{j_1} ... x_{j_k} -$$

$-m$-й элементарный симметрический многочлен. Из формулы Тейлора для функции $F_k(z)$ (см.,например, [7, с. 67]) получаем при $l < k$

$$f_k(t) = F_k(1) = f_1^k + f_1^{k-1} \mathbf{E} S_1(X_1, ..., X_k) +$$

$$+ f_1^{k-2} \mathbf{E} S_2(X_1, ..., X_k) + ... + f_1^{k-l} \mathbf{E} S_l(X_1, ..., X_k) + R_l, \tag{6}$$

где

$$R_l = R_l(X_1, ..., X_k) = \frac{1}{2\pi i} \int\limits_{|z|=\rho} \frac{F_k(z)\, dz}{(z-1)\, z^{l+1}}, \quad |\rho| > 1. \tag{7}$$

Пусть $g_l = \mathbf{E} X_1 ... X_l,\ l \ge 1,\ \beta^2 = \mathbf{E}|X_1|^2 = 1 - |f_1|^2$.

В дальнейшем будем обозначать через $\theta_i = \theta_i(t, n, ...),\ i = 1, 2, , ...$ - ограниченные величины, (т.е. $\sup\limits_{t,n,...} |\theta_i(t, n, ...)| < \infty$); если $g(t, n, ...) \le ch(t, n, ...)$, где $c \ge 0$ не зависит от $t, n, ...$, то будем писать $g \ll h$, а $g \asymp h$ будет обозначать, что $g \ll h$ и $h \ll g$.

Лемма 6. *1. Если последовательность $\{\xi_n\}$ удовлетворяет условию φ–перемешивания с коэффициентом перемешивания $\varphi(n)$, то*

a) $\mathbf{E} S_1(X_1, ..., X_k) = 0$;

b) $\mathbf{E} S_2(X_1, ..., X_k) = (k-1) g_2 + O(\beta^2 k^2 \varphi^{\frac{1}{2}}(p))$;

c) $\mathbf{E} S_3(X_1, ..., X_k) = (k-2) g_3 + O(\beta^2 k^3 \varphi^{\frac{1}{2}}(p))$;

d) $\mathbf{E} S_4(X_1, ..., X_k) = (k-3) g_4 + \frac{(k-3)(k-4)}{2} g_2^2 + O(\beta^2 k^4 \varphi^{\frac{1}{2}}(p))$.

Доказательство. a) $\mathbf{E} S_1(X_1, ..., X_k) = \mathbf{E} \sum\limits_{j=1}^{k} X_j = 0$.

b) В сумме $\mathbf{E} S_2(X_1, ..., X_k)$ имеется $k-1$ слагаемое вида $\mathbf{E} X_j X_{j+1} = g_2$ и $(k-1)(k-2)/2$ слагаемых вида $\mathbf{E} X_j X_l,\ l > j+1$, каждое из которых по лемме 3 не превосходит $2\varphi^{\frac{1}{2}}(p) \|X_j\|_2 \|X_l\|_2 == 2\beta^2 \varphi^{\frac{1}{2}}(p)$, что дает нам нужную оценку для $\mathbf{E} S_2(X_1, ..., X_k)$. Соотношение c) доказывается аналогично с учетом того, что $|X_j| \le 2$ при всех j .

d) В сумме $\mathbf{E} S_4(X_1, ..., X_k)$ имеется $k-3$ слагаемых вида $\mathbf{E} X_j X_{j+1} X_{j+2} X_{j+3} = g_4$ и $(k-3)(k-4)/2$ слагаемых вида $\mathbf{E} X_j X_{j+1} X_l X_{l+1}$, $l > j+2$, каждое из которых по лемме 3 равно $g_2^2 + O\left(\beta^2 \varphi^{\frac{1}{2}}(p)\right)$. Все прочие слагаемые имеют вид $\mathbf{E} X_i X_j X_k X_l$, где либо $j > i+1$, либо $l > k+1$; по лемме 3 такие слагаемые равны $O\left(\beta^2 \varphi^{\frac{1}{2}}(p)\right)$, а количество этих слагаемых равно $C_k^4 - (k-3) - (k-3)(k-4)/2 = O\left(k^4\right)$. Сказанное доказывает утверждение d).

Лемма 7. *Пусть* $n = kp$. *Если* $\mathbf{E}|\xi_1|^5 < \infty$, *и* $\sum\limits_{j=1}^{\infty} j\varphi^{\frac{1}{2}}(j) < \infty$, *то существует* $\varepsilon > 0$ *такое, что при* $|t| \le \varepsilon\sqrt{k}$

a)

$$f_1(t) = \exp\left\{-\frac{t^2}{2k} + \frac{t^4\gamma_4}{4!\sigma_1^4 nk} + \theta_1\left(\frac{t^4}{n^2} + \frac{|t|^5}{k^{\frac{5}{2}}}\right)\right\}. \tag{8}$$

b)

$$|g_l^*| \ll \frac{t^4}{n^2} + \frac{|t|^5}{k^{\frac{5}{2}}},\ g_l^* = f_1^{-l} g_l,\ l = 2, 3, 4. \tag{9}$$

Д о к а з а т е л ь с т в о. a) В силу леммы 5

$$\ln f_1(t) = \sum_{l=1}^{4} \frac{(it)^l \gamma_l(T_p)}{l!\sigma_n^l} + r_4(t) = -\frac{t^2}{2k} + \frac{t^4\gamma_4}{4!\sigma_1^4 nk} + \theta_2\frac{t^4}{n^2} + r_4(t), \tag{10}$$

где

$$r_4(t) = \frac{t^5}{5!}\left(\frac{d^5}{dt^5}\ln f_1(t)\right)_{t=c},\ c \in (0, t).$$

Производную $\left(\frac{d^5}{dt^5}\ln f_1(t)\right)_{t=c}$ можно представить в виде дроби со знаменателем $f_1^5(c)$, а слагаемые в числителе являются произведениями производных $f_1^{(l)}(c)$, $l = 0, 1, ..., 5$ суммарный порядок которых равен 5 (считаем $f_1^{(0)}(c) = f_1(c)$). Так как при $|t| \le \varepsilon\sqrt{k}$

$$|f_1(t) - 1| \le \frac{t^2\sigma_p^2}{2\sigma_n^2} \le \frac{\varepsilon^2}{2},$$

то при достаточно малых ε $|f_1(c)| \ge \frac{1}{2}$. Далее, с учетом леммы 4 получаем

$$|f_1^{(l)}(c)| \le \frac{\mathbf{E}|T_p|^l}{\sigma_n^l} \ll \left(\frac{\sigma_p}{\sigma_n}\right)^l = k^{-\frac{l}{2}}.$$

Отсюда следует, что $|r_4(t)| \ll \left(\frac{|t|}{\sqrt{k}}\right)^5$ и из (10) следует теперь (8).

b) Нетрудно подсчитать, что $g_2^* = f_1^{-2}f_2 - 1$, $g_3^* = \left(f_1^{-3}f_3 - 1\right) - 2g_2^* - g_2^{**}$, $g_4^* = \left(f_1^{-4}f_4 - 1\right) - 2g_3^* - 2g_3^{**} + 3g_2^* + 2g_2^{**} + g_2^{***}$, где

$$g_2^{**} = \mathbf{E}X_1X_3,\ g_2^{***} = \mathbf{E}X_1X_4,\ g_3^{**} = f_1^{-3}\mathbf{E}\exp\{it(Q_1 + Q_2 + Q_4)\} - 1.$$

Представив все входящие в эти выражения характеристические функции в виде (8) и воспользовавшись справедливым при любом комплексном z неравенством $|\exp\{z\} - 1| \le |z|\exp\{|z|\}$, последовательно оценим g_2, g_3, g_4; при $|t| \le \varepsilon\sqrt{k/l}$ получим

$$|g_l^*| \ll \left|\exp\left\{\theta_{3l}\left(\frac{t^4}{n^2} + \frac{|t|^5}{k^{\frac{5}{2}}}\right)\right\} - 1\right| \ll \frac{t^4}{n^2} + \frac{|t|^5}{k^{\frac{5}{2}}},\quad l = 2, 3, 4.$$

Лемма доказана.

Пусть $2q < p$. Обозначим

$$U_j = \sigma_n^{-1} \sum_{l=1}^{p-2q} \xi_{(j-1)p+q+l}, \quad V_j = \sigma_n^{-1} \sum_{l=1}^{q} (\xi_{(j-1)p+l} + \xi_{jp-q+l}),$$

$$u = u(t) = \mathbf{E}\exp\{itU_1\}, \quad v = v(t) = \mathbf{E}\exp\{itV_1\}.$$

Тогда $Q_j = U_j + V_j,\ X_j = Y_j + Z_j,\ j = 1, 2, ...,$ где

$$Y_j = (\exp\{itU_j\} - u)v, \quad Z_j = \exp\{itU_j\}(\exp\{itV_j\} - v) + uv - f_1,$$

так что

$$F_k(z) = \mathbf{E} \prod_{j=1}^{k} (f_1 + z(Y_j + Z_j)). \tag{11}$$

Лемма 8.

$$|F_k(z)| \le \left\{1 - \beta^2 + \rho^2\delta^2 + 8\rho^2\varphi(p)\right\}^{\frac{k}{2}} + 16\rho\varphi(q)B_k, \tag{12}$$

где $\delta^2 = 1 - |v|^2, \quad \rho = |z| \ge 1,$ *а*

$$B_k = \sum_{l=2}^{k-1} \| \prod_{j=1}^{l-1} (f_1 + zX_j)\|_\infty \left\{1 - \beta^2 + \rho^2\delta^2 + 8\rho^2\varphi(p)\right\}^{\frac{k-l-1}{2}} +$$

$$+ \| \prod_{j=1}^{k-1} (f_1 + zX_j)\|_\infty \left\{1 - \beta^2 + \rho^2\delta^2 + 8\rho^2\varphi(p)\right\}^{\frac{k-2}{2}}.$$

Доказательство.

Так как $\mathbf{E}Y_j = 0,\ |Y_j| \le 2,$ то с помощью леммы 3 при $1 < l < k$ получаем

$$|\mathbf{E} \prod_{j=1}^{l-1} (f_1 + zX_j) zY_l \prod_{j=l+1}^{k} (f_1 + zZ_j)| \le$$

$$\le |\mathbf{E} \prod_{j=1}^{l-1} (f_1 + zX_j) \mathbf{E} zY_l \prod_{j=l+1}^{k} (f_1 + zZ_j)| +$$

$$+ 4\rho\varphi(q) \| \prod_{j=1}^{l-1} (f_1 + zX_j)\|_\infty \| \prod_{j=l+1}^{k} (f_1 + zZ_j)\|_1 \le$$

$$\le 4\rho\varphi(q) |\mathbf{E} \prod_{j=1}^{l-1} (f_1 + zX_j)| \cdot \| \prod_{j=l+1}^{k} (f_1 + zZ_j)\|_1 +$$

$$+ 4\rho\varphi(q) \| \prod_{j=1}^{l-1} (f_1 + zX_j)\|_\infty \| \prod_{j=l+1}^{k} (f_1 + zZ_j)\|_1 \le$$

$$\leq 8\rho\varphi(q)\, b_l, \quad b_l = \| \prod_{j=1}^{l-1}(f_1 + zX_j)\|_\infty \, \| \prod_{j=l+1}^{k}(f_1 + zZ_j)\|_1. \tag{13}$$

Положим

$$b_1 = \| \prod_{j=2}^{k}(f_1 + zZ_j)\|_1, \quad b_k = \| \prod_{j=1}^{k-1}(f_1 + zX_j)\|_\infty.$$

Применяя последовательно соотношение (13), из (11) получаем

$$|F_k(z)| \leq |\mathbf{E}\prod_{j=1}^{k-1}(f_1 + zX_j)zY_k| + |\mathbf{E}\prod_{j=1}^{k-1}(f_1 + zX_j)(f_1 + zZ_k)| \leq$$

$$\leq |\mathbf{E}\prod_{j=1}^{k-1}(f_1 + zX_j)(f_1 + zZ_k)| + 8\rho\varphi(q)b_k \leq ... \leq$$

$$\leq |\mathbf{E}\prod_{j=1}^{k}(f_1 + zZ_j)| + 8\rho\varphi(q)\sum_{l=1}^{k} b_l. \tag{14}$$

Имеем

$$\mathbf{E}\prod_{j=1}^{r}|f_1 + zZ_j| \leq \left(\mathbf{E}\prod_{j\geq 1}{}'|f_1 + zZ_j|^2 \cdot \mathbf{E}\prod_{j\geq 2}{}''|f_1 + zZ_j|^2\right)^{\frac{1}{2}}, \tag{15}$$

где $\prod'$ и $\prod''$ обозначают произведения по нечетным и четным индексам соответственно. Так как $\mathbf{E}Z_1 = 0$, то

$$\mathbf{E}|Z_1|^2 \leq \mathbf{E}|Z_1 - uv + f_1|^2 = \mathbf{E}|\exp\{itV_1\} - v|^2 = 1 - |v|^2 = \delta^2.$$

Применяя последовательно соотношение (4), и учитывая, что при $\rho \geq 1$, $a = |f_1|^2 - 2\mathbf{Re}zf_1(uv - f_1)$ имеет место $||f_1 + zZ_j|^2 - a| \leq 4\rho^2$, получаем

$$\mathbf{E}\prod_{j\geq 1}{}'|f_1 + zZ_j|^2 \leq \left(\mathbf{E}|f_1 + zZ_1|^2 + 8\rho^2\varphi(p)\right)\mathbf{E}\prod_{j\geq 3}{}'|f_1 + zZ_j|^2 \leq ...$$

$$\leq \left\{1 - \beta^2 + \rho^2\delta^2 + 8\rho^2\varphi(p)\right\}^{\left[\frac{r+1}{2}\right]}. \tag{16}$$

Аналогично

$$\mathbf{E}\prod_{j\geq 2}{}''|f_1 + z_j|^2 \leq \left\{1 - \beta^2 + \rho^2\delta^2 + 8\rho^2\varphi(p)\right\}^{\left[\frac{r}{2}\right]}. \tag{17}$$

Из (15), (16) и (17) следует

$$\mathbf{E}\prod_{j=1}^{r}|f_1 + zZ_j| \leq \left\{1 - \beta^2 + \rho^2\delta^2 + 8\rho^2\varphi(p)\right\}^{\frac{r}{2}}. \tag{18}$$

Из (14) и (18) вытекает утверждение леммы.

3. Основные результаты

Теорема 1. *Пусть* $\{\xi_n\}$ *– стационарная последовательность с симметричным распределением и экспоненциально быстрым* φ*–перемешиванием и пусть* $\mathbf{E}|\xi_1|^5 < \infty$ *и выполнено условие (1). Тогда*

$$F_n(x) = \Phi(x) + \frac{\gamma_4(3x - x^3)}{4!\sqrt{2\pi}\sigma_1^4 n} \exp\left\{-\frac{x^2}{2}\right\} + O\left(n^{-\frac{6}{5}}\right).$$

Для доказательства потребуется еще несколько лемм.

Лемма 9. *При выполнении условий теоремы 1 существует* $\varepsilon > 0$ *такое, что*

$$f_n^*(t) = \mathbf{E}\exp\left\{\frac{itT_n}{\sigma_n}\right\} = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_4\left(\frac{t^4 + |t|^5}{n^{\frac{6}{5}}}\right)\right\}, \tag{19}$$

где $|t| \le \varepsilon\sqrt{k}, \quad k = k(n) = [n^a], \quad a = 1 - \frac{1}{\sqrt[6]{5}} \approx 0,235.$

Д о к а з а т е л ь с т в о.

Если $|t| \le n^{-2}$, то, положив в (8) $p = n,\ k = 1$ получим

$$f_k(t) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_5\frac{t^4}{n^2}\right\},$$

откуда следует (19).

Пусть теперь $n^{-2} \le |t| \le \varepsilon\sqrt{k}$. Положим

$$q = \left[n^{0,5-a}\right], \quad p = \left[n^{1-a}\right], \quad n = kp.$$

Аналогично (10) нетрудно получить

$$\ln|v(t)|^2 = -\frac{2t^2 q}{n} + \theta_6\frac{t^4 q^2}{n^2} \sim -\frac{t^2 q}{n} \sim |v(t)|^2 - 1, \quad n \to \infty.$$

откуда при достаточно больших n выводим

$$\delta^2 = 1 - |v|^2 \le \frac{3t^2 q}{n}, \quad \delta^2 \ge \frac{t^2 q}{n} \ge \frac{q}{n^5}, \tag{20}$$

$$\beta^2 = \mathbf{E}\left|\exp\{itX_1\} - f_1\right|^2 \le \mathbf{E}\left|\exp\{itX_1\} - 1\right|^2 \le \frac{t^2\sigma_p^2}{\sigma_n^2} \le \frac{1}{2}. \tag{21}$$

Положив в (6) и (7) $l = 4$, с помощью леммы 6 получаем

$$f_k(t) = f_1^k(t)\left\{1 + (k-1)g_2^* + (k-2)g_3^* + (k-3)g_4^* + \right.$$

$$\left. + \frac{(k-3)(k-4)}{2}(g_2^*)^2 + \theta_7\beta^2 k^4\varphi^{\frac{1}{2}}(p)\right\} + R_4, \tag{22}$$

$$|R_4| = \left| \frac{1}{2\pi i} \int\limits_{|z|=\rho} \frac{F_k(z)\,dz}{(z-1)\,z^5} \right| \le (\rho-1)^{-1}\rho^{-5} \sup_{|z|=\rho} |F_k(z)|, \quad |\rho| > 1. \tag{23}$$

Положим в (12) $\rho^2 = \dfrac{1-\beta^2}{k\delta^2}$. При достаточно малых ε и достаточно больших n с помощью (20), (21), (23) и леммы 8 получаем

$$\rho^{-2} = (1-\beta^2)^{-1}k\delta^2 \le 6\frac{kqt^2}{n} \le 6\varepsilon^2k^2qn^{-1} \le \frac{1}{4}, \quad \rho^2 \le \frac{n^5}{qk} \le \frac{n^5}{16}.$$

$$|R_4| \le 90 f_1^k \left(\frac{kq}{n}\right)^{\frac{5}{2}} |t|^5 \left(1 + \frac{1}{k} + n^5\varphi(p)\right)^{\frac{k}{2}} + 16t^4n^9\varphi(q)B_k \ll$$

$$\ll f_1^k \left(\frac{|t|}{\sqrt{n}}\right)^5 + t^4n^9\varphi(q)B_k, \tag{24}$$

где

$$B_k \le k(1+2\rho)^k \left(1 + \frac{1}{k} + n^5\varphi(p)\right)^{\frac{k}{2}} \ll k(1+n^3)^k. \tag{25}$$

По условию $\varphi(n) \le \exp\{-\alpha n\}$, и так как $k\ln n = o(q)$, то из (25) следует

$$n^9\varphi(q)B_k \ll \exp\left\{9\ln n + k\ln(1+n^3) - \alpha q\right\} \sim \exp\{-\alpha q\},$$

и соотношение (24) можно переписать так:

$$|R_4| \ll f_1^k \left(\frac{|t|}{\sqrt{n}}\right)^5 + t^4 \exp\{-\alpha q\}. \tag{26}$$

Из (22) с помощью (9), (21), и (26) получаем

$$f_k(t) = f_1^k(t) \left\{1 + \theta_8 \left(\frac{kt^4}{n^2} + \frac{|t|^5}{k^{\frac{3}{2}}}\right)\right\} + \theta_9 t^4 \exp\{-\alpha q\}, \tag{27}$$

Далее, из (8) следует

$$f_1^k(t) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_1 \left(\frac{kt^4}{n^2} + \frac{|t|^5}{k^{\frac{3}{2}}}\right)\right\}, \tag{28}$$

откуда с помощью (27) выводим

$$f_k(t) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_{10} \left(\frac{kt^4}{n^2} + \frac{|t|^5}{k^{\frac{3}{2}}}\right)\right\} + \theta_9 t^4 \exp\{-\alpha q\}. \tag{29}$$

Нетрудно видеть, что если $|t| \le \varepsilon\sqrt{k}$, то при достаточно малых $\varepsilon > 0$

$$\exp\{-\alpha q\} = o\left(n^{-2} \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + |\theta_{10}| \left(\frac{kt^4}{n^2} + \frac{|t|^5}{k^{\frac{3}{2}}}\right)\right\}\right),$$

так что (29) можно переписать так:

$$f_k(t) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_{11}\left(\frac{kt^4}{n^2} + \frac{|t|^5}{k^{\frac{3}{2}}}\right)\right\}. \tag{30}$$

Будем считать (30) первой итерацией для $f_k(t)$. Обозначим $k_1 = k_1(n) = k(p) = [p^a]$. Заменив в (30) n на p, k на k_1, получим

$$f_1\left(t\frac{\sigma_n}{\sigma_p}\right) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 p} + \theta_{12}\left(\frac{k_1 t^4}{p^2} + \frac{|t|^5}{k_1^{\frac{3}{2}}}\right)\right\}, \quad |t| \le \varepsilon\sqrt{k_1}.$$

откуда

$$f_1(t) = \exp\left\{-\frac{t^2}{2k} + \frac{t^4\gamma_4}{4!\sigma_1^4 nk} + \theta_{13}\left(\frac{k_1 t^4}{n^2} + \frac{|t|^5}{k^{\frac{5}{2}}k_1^{\frac{3}{2}}}\right)\right\}, \quad |t| \le \varepsilon\sqrt{kk_1}$$

Отсюда, повторив доказательство леммы 7, можно получить

$$|g_l^*| \ll \frac{k_1 t^4}{n^2} + \frac{|t|^5}{k^{\frac{5}{2}}k_1^{\frac{3}{2}}}, \quad l = 2, 3, 4.$$

Если теперь провести приведенные выше рассуждения настоящей леммы с использованием двух последних соотношений вместо (8) и (9), то можно получить вторую итерацию для $f_k(t)$:

$$f_k(t) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_{14}\left(\frac{kk_1 t^4}{n^2} + \frac{|t|^5}{(kk_1)^{\frac{3}{2}}}\right)\right\}, \quad |t| \le \varepsilon\sqrt{k},$$

и т.д. После шестой итерации получим

$$f_k(t) = \exp\left\{-\frac{t^2}{2} + \frac{t^4\gamma_4}{4!\sigma_1^4 n} + \theta_{15}\left(\frac{Kt^4}{n^2} + \frac{|t|^5}{K^{\frac{3}{2}}}\right)\right\}, \quad |t| \le \varepsilon\sqrt{k}, \tag{31}$$

где $K = kk_1k_2k_3k_4k_5 \sim n^A$, $A = a\sum\limits_{l=0}^{5}(1-a)^l = 1-(1-a)^6 = \frac{4}{5}$. Из (31) следует теперь утверждение леммы в случае, когда $n = kp$.

Пусть теперь $n = kp + r,\ 0 \le r < n$. Введем $Q_j,\ j = 1, ..., k$ по формулам (5) и положим

$$\tilde{Q}_{k+1} = \sigma_n{}^{-1}\sum_{l=1}^{r}\xi_{kp+l},\ \tilde{f}_1(t) = \mathbf{E}\exp\{it\tilde{Q}_{k+1}\},\ \tilde{X}_{k+1} = \exp\left\{it\tilde{Q}_{k+1}\right\} - \tilde{f}_1.$$

Необходимые изменения в доказательстве равенства (19) достаточно прозрачны: главное - для $f_1^k\tilde{f}_1$ сохраняется то же представление, что и для f_1^k в (28), а новые (по сравнению с $\mathbf{E}S_l(X_1, ..., X_k)$) слагаемые в $\mathbf{E}S_l(X_1, ..., X_k, \tilde{X}_{k+1})$ и «лишний» (по сравнению с $F_k(z)$) сомножитель в $\mathbf{E}\prod\limits_{j=1}^{k}(f_1 + zX_j)(\tilde{f}_1 + z\tilde{X}_{k+1})$ оцениваются в общем так же, как и старые. Доказательство остается принципиально тем же, хотя и более громоздким.

Лемма 10. *При* $\varepsilon\sqrt{k} \le |t| \le \varepsilon\sqrt{n},\ \varepsilon > 0.$

$$|f_n^*(t)| \le \exp\left\{-bt^2\right\}, \quad b > 0$$

Д о к а з а т е л ь с т в о. Пусть $kp \le n < (k+1)p$. Используя замечание 2 и лемму 1, аналогично (18) получаем

$$|f_n^*(t)| \le \left|\mathbf{E}\prod_{j=1}^{k}\cos\left(tQ_1\right)\right| \le \left(\mathbf{E}\cos^2\left(tQ_j\right) + 2\varphi(p)\right)^{\frac{k}{2}} \le$$

$$\le \left\{\frac{1+f_1(2t)}{2} + 2\varphi(p)\right\}^{\frac{k}{2}}. \qquad (32)$$

Пусть $k = k(n) = \left[\frac{t^2}{N}\right],\ p = \left[\frac{n}{k}\right]$. Тогда $p = p(n) \sim \frac{nN}{t^2} \ge \frac{N}{\varepsilon^2}$, В силу леммы 9

$$f_1(2t) \asymp \exp\left\{-\frac{2t^2}{k}\right\} \sim \exp\left\{-2N\right\}$$

так что $N > 0$ можно подобрать так, чтобы $2\varphi(p) < \frac{1}{8},\quad f_1(2t) < \frac{1}{4}$, и из (32) следует теперь

$$|f_n^*(t)| \le \left(\frac{3}{4}\right)^{\frac{k}{2}} \sim \exp\left\{-\frac{(\ln 4 - \ln 3)}{2N}t^2\right\}.$$

Лемма 11. *При* $|t| \ge \varepsilon\sqrt{n}$

$$|f_n^*(t)| \le \mu^n, \quad 0 < \mu < 1,$$

Д о к а з а т е л ь с т в о. Если имеет место (1), то для любого $\varepsilon > 0$ найдется $\tau < 1$ такое, что $|\mathbf{E}\exp\left\{it\xi_1\right\}| = |\mathbf{E}\cos\left\{it\xi_1\right\}| \le \tau, \quad |t| \ge \frac{\varepsilon}{2}$ (см., например, [2, с.22]). Выберем натуральное p так, чтобы
$2\varphi(p) < (1-\tau)/8$. Тогда в силу лемм 3 и 1

$$|f_n^*(t)| = \left|\mathbf{E}\prod_{j=1}^{n}\cos\left\{it\sigma_n^{-1}\xi_j\right\}\right| \le \mathbf{E}\prod_{j=1}^{\left[\frac{n}{p}\right]}\left|\cos\left\{it\sigma_n^{-1}\xi_{jp}\right\}\right| \le$$

$$\le \left(\mathbf{E}\left|\cos\left(t\sigma_n^{-1}\xi_1\right)\right| + 2\varphi(p)\right)^{\left[\frac{n}{p}\right|} \le \left(\sqrt{\frac{1+\tau}{2}} + 2\varphi(p)\right)^{\left[\frac{n}{p}\right|} \le$$

$$\le \left(1 - \frac{1-\tau}{8}\right)^{\left[\frac{n}{p}\right|}, \quad |t| \ge \varepsilon\sigma_n|t|.$$

Из последнего соотношения следует утверждение леммы.

Д о к а з а т е л ь с т в о т е о р е м ы 1.

Воспользуемся следующим вариантом классического неравенства Эссеена:

$$|F(x) - G(x)| \leq \frac{1}{\pi} \int\limits_{-T}^{T} \left| \frac{f(t) - g(t)}{t} \right| + \frac{24m}{\pi T}, \tag{33}$$

где F - функция распределения, f – соответствующая ей характеристическая функция, G – функция ограниченной вариации, g – ее преобразование Фурье, $F(x) - G(x) \to 0,\ x \to \pm\infty,\quad g(0) = 1,\ f'(0) = g'(0) = 0,\ |G'(x)| \leq m$ (см.,например, [9, с.603]).

Положим в (33)

$$F(x) = F_n(x), \quad G(x) = \Phi(x) + \frac{\gamma_4(3x - x^3)}{4!\sqrt{2\pi}\sigma_1^4 n} \exp\left\{-\frac{x^2}{2}\right\}.$$

Тогда

$$f(t) = f_n^*(t), \quad g(t) = \exp\left\{-\frac{t^2}{2}\right\}\left(1 + \frac{\gamma_4 t^4}{4!\sigma_1^4 n}\right)$$

(см., например [2, с.180]).

Обозначим $\Gamma_n(t) = \dfrac{\gamma_4 t^4}{4!\sigma_1^4 n}$. С помощью соотношения (19) получаем при $|t| \leq \varepsilon\sqrt{k},\quad k = [n^a],\quad a = 1 - \dfrac{1}{\sqrt[6]{5}}$

$$|f(t) - g(t)| = \exp\left\{-\frac{t^2}{2}\right\} \left| \exp\left\{\Gamma_n(t) + \theta_4\left(\frac{t^4 + |t|^5}{n^{\frac{6}{5}}}\right)\right\} - 1 - \Gamma_n(t) \right|.$$

Так как $\Gamma_n(t) \to 0,\ \dfrac{t^4 + |t|^5}{n^{\frac{6}{5}}} \to 0$ при $|t| \leq \varepsilon\sqrt{k},\ n \to \infty$, то, воспользовавшись справедливым при любых комплексных u и v неравенством

$$|\exp\{u + v\} - 1 - v| \leq \left(|u| + \frac{1}{2}|v|^2\right) \exp\{\max(|u|, |v|)\},$$

получаем

$$\left| \exp\left\{\Gamma_n(t) + \theta_4\left(\frac{t^4 + |t|^5}{n^{\frac{6}{5}}}\right)\right\} - 1 - \Gamma_n(t) \right| \ll$$

$$\ll \frac{t^4 + |t|^5}{n^{\frac{6}{5}}} + \Gamma_n^2(t) \ll \frac{t^4 + |t|^5}{n^{\frac{6}{5}}}.$$

Следовательно

$$|f(t) - g(t)| \ll \frac{t^4 + |t|^5}{n^{\frac{6}{5}}} \exp\left\{-\frac{t^2}{2}\right\}, \quad |t| \leq \varepsilon\sqrt{k}. \tag{34}$$

Положив теперь в (33) $T = n^{\frac{6}{5}}$ и воспользовавшись для оценки $|f(t) - g(t)|$ при $|t| \leq \varepsilon\sqrt{k}$ соотношением (34), а для оценки $|f(t)| = |f_n^*(t)|$ при

$\varepsilon\sqrt{k} \le |t| \le \varepsilon\sqrt{n}$ – леммой 10 и при $\varepsilon\sqrt{n} \le |t| \le n^{\frac{6}{5}}$ – леммой 11, получим утверждение теоремы.

Литература

1. Rio, E. Sur le theoreme de Berry-Esseen pour les suites faiblement dependantes. (French) / E. Rio // J.Probab. Theory Relat. Fields. – 1996. – V.104. №2. – P.255-282.
2. Петров, В.В. Предельные теоремы для сумм независимых случайных величин / В.В. Петров. – М.: Наука, 1987.
3. Bradley, R. On the φ-mixing condition for stationary random sequences / R. Bradley // Duke Mathematical Journal. – 1980. – V.47. №2. – P.421–433.
4. Ибрагимов, И.А. Независимые и стационарно связанные величины / И.А. Ибрагимов, Ю.В. Линник. – М.: Наука, 1965.
5. Биллингсли, П. Сходимость вероятностных мер / П. Биллингсли. – М: Наука, 1977, 351 с.
6. Peligrad, M. The convergence of moments in the central limit theorem for ρ-mixing sequences of random variables / M. Peligrad // Pros. of the American Math. Soc. – 1987. – V.101. №1. – P.142-148.
7. Лаврентьев, М.А. Методы теории функций комплексного переменного / М.А. Лаврентьев, Б.В. Шабат. – М.: Наука, 1973.
8. Лоэв М. Теория вероятностей / М. Лоэв. – М.: ИЛ, 1962.
9. Феллер, В. Введение в теорию вероятностей и ее приложения. Т.2. / В. Феллер. – М.: Мир, 1984.

Математические структуры и моделирование
2009, вып. 19, с. 19–27

УДК 517.91

МЕТОД ОБРАТНОЙ ФУНКЦИИ ДЛЯ РЕШЕНИЯ АВТОНОМНОГО ОБЫКНОВЕННОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ: ОДНОМЕРНЫЙ СЛУЧАЙ

В.В. Коробицын, Ю.В. Фролова

A novel method for numerical solving autonomous ordinary differential equation $dx/dt = f(x)$ is suggested. The solution is found as a function $t(x)$ for that the incrementation is independent from previous points of solution trajectory. This fact is used for parallel realization of computational method. The numerical experiments show the applicability of suggested method for area with rapid acceleration of right-hand function.

Численные методы решения обыкновенных дифференциальных уравнений появились довольно давно. Еще Ньютон и Эйлер предложили методы поиска приближенного решения. Классические работы Рунге, Кутты, Адамса посвящены построению одношаговых и многошаговых методов решения ОДУ. Казалось бы всё уже разработано и внедрено, однако интерес к разработке новых и модификации существующих методов не ослабевает. С одной стороны, новые задачи требуют новых методов. Так в последнее время активно развиваются методы решения жестких, дифференциально-алгебраических и разрывных задач [1]–[4], [8]– [11]. А с другой стороны, развитие компьютерной техники требует разработки эффективных методов решения на многоядерных, многопроцессорных и распределенных компьютерных системах. В последнее время стали появляться работы связанные с параллельной реализацией методов решения ОДУ [5]–[7], [12].

Задача данной статьи заключается в представлении нового метода решения обыкновенных дифференциальных уравнений с высокой степенью распараллеливания вычислений. Решение вычисляется независимо на отдельных промежутках по пространственной переменной x, где определяется время τ, необходимое для преодоления промежутка.

Представленный метод применим для решения одного автономного нелинейного обыкновенного дифференциального уравнения первого порядка. А также высказана идея о распространении этого метода на автономные системы ОДУ.

Омский государственный университет.
E-mail: `korobits@rambler.ru`

1. Аппроксимация решения автономного ОДУ

Рассмотрим задачу
$$\frac{dx}{dt} = f(x),\ x(t_0) = x_0,\ t > t_0. \tag{1}$$
Необходимо найти решение $x(t) : \mathbb{R} \to \mathbb{R}$ задачи Коши (1), учитывая, что функция правой части f является непрерывной по переменной x и не зависит от t в явном виде.

Решение $x(t)$ можно разложить в ряд Тейлора вблизи точки t_0
$$x(t_0+\tau) = x(t_0) + \frac{dx}{dt}(t_0)\cdot\tau + \frac{d^2x}{dt^2}(t_0)\cdot\frac{\tau^2}{2!} + \dots, \tag{2}$$
Частичную сумму ряда (2) обозначим за $P_k(t_0,\tau)$
$$P_k(t_0,\tau) := \frac{dx}{dt}(t_0)\cdot\tau + \frac{d^2x}{dt^2}(t_0)\cdot\frac{\tau^2}{2!} + \dots + \frac{d^kx}{dt^k}(t_0)\cdot\frac{\tau^k}{k!},$$
тогда
$$x(t_0+\tau) - x(t_0) = P_k(t_0,\tau) + O(\tau^{k+1}).$$
Введем обозначения
$$f_0 = f(x_0),\ f_0' = \frac{df}{dx}(x_0),\ f_0'' = \frac{d^2f}{dx^2}(x_0)$$
и приведем выражения для P_k, $k = 0, 1, 2, \dots$:
$$P_0(t_0,\tau) = 0,$$
$$P_1(t_0,\tau) = P_0(t_0,\tau) + \frac{dx}{dt}(t_0)\cdot\tau = f_0\tau,$$
$$P_2(t_0,\tau) = P_1(t_0,\tau) + \frac{d^2x}{dt^2}(t_0)\cdot\frac{\tau^2}{2!} = f_0\tau + f_0'f_0\frac{\tau^2}{2},$$
$$P_3(t_0,\tau) = P_2(t_0,\tau) + \frac{d^3x}{dt^3}(t_0)\cdot\frac{\tau^3}{3!} = f_0\tau + f_0'f_0\frac{\tau^2}{2} + \left(f_0''f_0^2 + f_0'^2f_0\right)\frac{\tau^3}{6}, \cdots$$

В отличие от традиционного подхода, когда по заданному шагу τ находят значение решения $x(t_0+\tau)$, мы зададим малое приращение d_x по переменной x и будем искать τ такое, что
$$x(t_0+\tau) - x(t_0) = d_x.$$
В этом случае, необходимо решить алгебраическое уравнение
$$P_k(t_0,\tau) = d_x. \tag{3}$$
Функция $P_k(t_0,\tau)$ является многочленом от τ степени k, поэтому уранение (3) будет иметь k корней. Необходимо найти наименьший действительный положительный корень, который будет давать значение приращения по переменной t, необходимое для достижения точки $x_0 + d_x$.

Отметим, что решение уравнения (1) перейдет из точки x_0 в точку $x_0 + d_x$ при $d_x > 0$, только когда $f(x) > 0$ для всех $x \in [x_0; x_0 + d_x]$. Однако, если в некоторой точке x_1 этого интервала функция $f(x_1) = 0$, то это означает, что в этой точке уравнение (1) имеет особую точку. Поэтому решение не достигнет точки x_0+d_x. Если же $f(x_0) < 0$, то изучается интервал $[x_0; x_0-d_x]$ аналогичным образом.

2. Формулы для вычисления τ

Далее будем считать, что $f(x) > 0$ для всех $x \in [x_0; x_0+d_x]$. Введем обозначения

$$f_1 = f(x_0 + d_x),\ \ f_{-1} = f(x_0 - d_x).$$

Тогда значения производных от функции f в точке x_0 можно оценить формулами

$$f_0' \approx f_0'(d_x) := \frac{f_1 - f_{-1}}{2d_x}, \quad f_0'' \approx f_0''(d_x) := \frac{f_1 - 2f_0 + f_{-1}}{d_x^2}. \tag{4}$$

Оба приближения имеют погрешность порядка $O(d_x^3)$.

Найдем аналитические выражения для корней уравнения (3) при разных k.

При $k = 1$ получаем уравнение $f_0\tau = d_x$. Корень этого линейного уравнения определяется выражением $\tau = \frac{d_x}{f_0}$. Погрешность решения, соответствующего найденному τ будет порядка $O(d_x^2)$.

При $k = 2$ получаем квадратное уравнение относительно τ

$$\frac{f_0' f_0}{2}\tau^2 + f_0\tau - d_x = 0,$$

корни которого определяются выражениями

$$D = f_0^2 + 2f_0 f_0' d_x, \qquad \tau_{1,2} = \frac{-f_0 \pm \sqrt{D}}{f_0 f_0'} = \frac{1}{f_0'}\left(-1 \pm \sqrt{1 + 2d_x\frac{f_0'}{f_0}}\right).$$

Если $f_0' > 0$, то положительный действительный корень будет

$$\tau = \frac{1}{f_0'}\left(\sqrt{1 + 2d_x\frac{f_0'}{f_0}} - 1\right). \tag{5}$$

Если $f_0' < 0$, то для существования действительного корня необходимо удовлетворение условия $1 + 2d_x f_0'/f_0 \geq 0$, что будет выполнено при $d_x \leq -f_0/(2f_0')$. Учитывая выражение для f_0' из (4) получаем $d_x \leq -f_0 d_x/(f_1 - f_{-1})$. Отсюда $-f_0/(f_1 - f_{-1}) \geq 1$. Следовательно, d_x нужно выбрать таким, чтобы было выполнено условие $f_1 - f_{-1} \geq -f_0$. Тогда минимальное τ для достижения $x_0 + d_x$ соответствует (5).

При $k \geq 3$ получаем алгебраическое полиномиальное уравнение относительно τ степени k. Это уравнение можно решать численно. При этом необходимо взять наименьший положительный корень τ, соответствующий первому пересечению кривой полинома с точкой $x + d_x$.

Однако мы предлагаем повысить точность вычисления τ, не повышая степени полинома. Для этого используем разложение функции решения в точке $x_1 = x(t_1)$, $t_1 = t_0 + \tau$, вычисляя значение функции решения в точке $x_0 = x_1 - d_x$, получаем

$$x(t_0) = x(t_1) - \frac{dx}{dt}(t_1) \cdot \tau + \frac{d^2x}{dt^2}(t_1) \cdot \frac{\tau^2}{2!} + \dots + (-1)^k \frac{d^kx}{dt^k}(t_1) \cdot \frac{\tau^k}{k!} + \dots,$$

Тогда $x(t_0) - x(t_1) = P_k(t_1, -\tau) + O(\tau^{k+1})$. Учитывая выражение (3), получаем два соотношения $x_1 - x_0 = P_k(t_0, \tau)$ и $x_0 - x_1 = P_k(t_1, -\tau)$. Вычитая эти соотношения, получаем

$$2d_x = P_k(t_0, \tau) - P_k(t_1, -\tau). \tag{6}$$

Введем обозначения

$$f_2 = f(x_0 + 2d_x), \quad f_1' = \frac{df}{dx}(x_1) \approx \frac{f_2 - f_0}{2d_x}, \quad f_1'' = \frac{d^2f}{dx^2}(x_1) \approx \frac{f_2 - 2f_1 + f_0}{d_x^2}. \tag{7}$$

Вычислим выражение для τ как корень уравнения (6) для разных значений параметра k.

При $k = 1$ получаем линейное уравнение $2d_x = f_0\tau - f_1 \cdot (-\tau)$. Решение этого уравнения $\tau = 2d_x/(f_0 + f_1)$. Погрешность определения величины τ будет порядка $O(d_x^2)$.

При $k = 2$ получаем квадратное уравнение

$$\left(f_0'f_0 - f_1'f_1\right)\frac{\tau^2}{2} + (f_0 + f_1)\tau - 2d_x = 0. \tag{8}$$

При $k \geq 3$ получаем полиномиальное уравнение относительно τ, наименьший положительный вещественный корень которого можно вычислять численно.

Далее мы найдем аналитическое выражение для наименьшего положительного вещественного корня уравнения (8) и оценим погрешность его вычисления при аппроксимации производных функции f, входящих в уравнение, конечными разностями вида (7).

Выражение для коэффициента уравнения при квадрате можно вычислить

$$f_0'f_0 - f_1'f_1 = \frac{f_1 - f_{-1}}{2d_x} f_0 - \frac{f_2 - f_0}{2d_x} f_1 = \frac{-f_{-1}f_0 + 2f_0f_1 - f_1f_2}{2d_x}.$$

Найдем корни уравнения (8)

$$D = \left(f_0 + f_1\right)^2 - 4\left(f_0'f_0 - f_1'f_1\right)\left(-2d_x\right) = \left(f_0 + f_1\right)^2 + 4\left(-f_{-1}f_0 + 2f_0f_1 - f_1f_2\right),$$

$$\tau_{1,2} = \frac{-\left(f_0 + f_1\right) \pm \sqrt{D}}{2(f_0'f_0 - f_1'f_1)} = d_x \frac{\left(f_0 + f_1\right) \mp \sqrt{\left(f_0 + f_1\right)^2 - 4\left(f_{-1}f_0 - 2f_0f_1 + f_1f_2\right)}}{\left(f_{-1}f_0 - 2f_0f_1 + f_1f_2\right)}.$$

Введем обозначение

$$F = \frac{f_0 + f_1}{f_{-1}f_0 - 2f_0f_1 + f_1f_2}$$

и перепишем выражение для корней уравнения

$$\tau_{1,2} = d_xF\left(1 \pm \sqrt{1 - \frac{4}{(f_0 + f_1)F}}\right).$$

Для существования вещественных корней уравнения необходимо выполнение условия $\frac{4}{\left(f_0+f_1\right)F} \le 1$, это будет выполнено в одном из случаев $(f_0 + f_1)F < 0$ или $(f_0 + f_1)F \ge 4$.

Наименьший положительный вещественный корень всегда определяется выражением

$$\tau = d_xF\left(1 - \sqrt{1 - \frac{4}{(f_0 + f_1)F}}\right). \tag{9}$$

Причем, если произведение $d_xF > 0$, то должно быть $\left(f_0 + f_1\right)F \ge 4$, а если $d_xF < 0$, то должно быть $\left(f_0 + f_1\right)F < 0$.

Каким должен быть d_x? Знак величины d_x определяется знаком функции f. Если $f_0 > 0$, то должно быть $d_x > 0$. Длина шага d_x будет зависеть от оценки погрешности.

Оценим погрешность найденного корня (9). При вычислении корня производные функции f были заменены конечными разностями (7), погрешность вычисления которых порядка $O(d_x^2)$. Следовательно, погрешность вычисления F будет порядка $O(d_x^3)$, а погрешность определения τ — порядка $O(d_x^4)$.

3. Практическая оценка погрешности

Для практической оценки погрешности вычисления значения τ проведем вычисления с двумя шагами длиной d_x и одним шагом длиной $2d_x$, тогда

$$t(x_0 + 2d_x) - t_0 = \tau_1 + \tau_2 + C \cdot (d_x)^p,$$

$$t(x_0 + 2d_x) - t_0 = \tau_3 + C \cdot (2d_x)^p,$$

где τ_1, τ_2 — два значения, вычисленные при шаге d_x, τ_3 — одно значение, соответствующее шагу $2d_x$. Вычитая первое выражение из второго, получим

$$\tau_1 + \tau_2 - \tau_3 = C \cdot (d_x)^p \cdot (2^p - 1).$$

Тогда погрешность вычисления τ с шагом d_x будет

$$Err_{d_x} = C \cdot (d_x)^p = \frac{\tau_1 + \tau_2 - \tau_3}{2^p - 1}.$$

4. Численные эксперименты

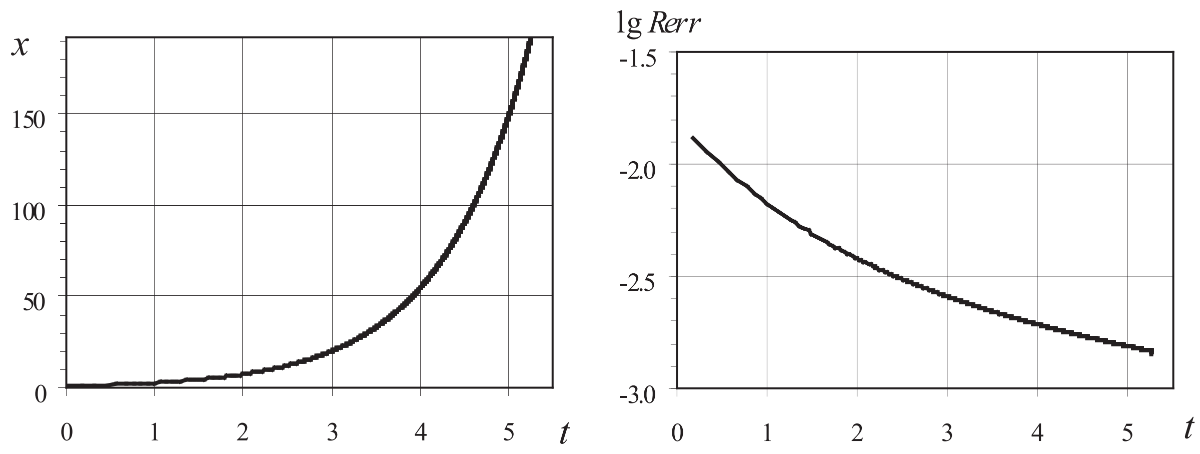
Точность предложенного метода нахождения численного решения автономного ОДУ оценим на двух примерах.

Первое уравнение — линейное:

$$\frac{dx}{dt} = x,\ x(0) = x_0,\ t > 0. \tag{10}$$

Аналитическое решение этого уравнения определяется формулой $x(t) = x_0 e^t$. Откуда можно выразить функцию $t(x) = \ln(x/x_0)$.

Для численного эксперимента использовалась формула (9). Были заданы параметры: $x_0 = 1$, $d_x = 0.19$. График решения уравнения (рис. 1-а) показывает экспоненциальный рост $x(t)$. Порядок относительной погрешности численного решения уменьшается от -1.833 до -2.845 (рис. 1-б). Судя по данному графику, можно сделать вывод, что предложенная схема вычисления решения лучше работает в условиях больших значений функции правой части уравнения.



а) б)

Рис. 1. а) График решения системы (10), б) Порядок относительной погрешности численного решения

Второе рассматриваемое уравнение — модель Верхюльста-Пёрла:

$$\frac{dx}{dt} = k(a-x)x,\ x(0) = x_0,\ t > 0. \tag{11}$$

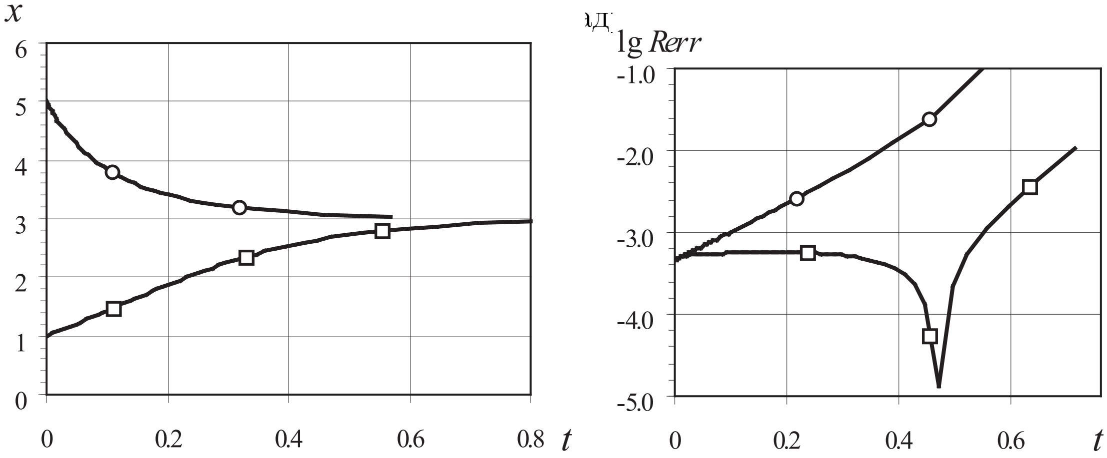
Аналитическим решением уравнения является логистическая кривая, определяемая формулами:

$$x(t) = \frac{ae^{ka(t-C_1)}}{1+e^{ka(t-C_1)}} \text{ при } 0 < x < a, \quad x(t) = \frac{ae^{ka(t-C_2)}}{1-e^{ka(t-C_2)}} \text{ при } x > a.$$

Значения констант определяются исходя из начальных данных

$$C_1 = \frac{-\ln\frac{x_0}{a-x_0}}{ka},\ C_2 = \frac{-\ln\frac{x_0}{x_0-a}}{ka}.$$

Численное решение построено при заданных параметрах: $a = 3$, $k = 2$, $d_x = 0.04$. Начальные данные для построения двух траекторий решения: $x_0^1 = 1$, $x_0^2 = 5$.



а) б)

Рис. 2. а) График решения системы (11), б) Порядок относительной погрешности численного решения; линии с квадратиками для $0 < x < a$, линии с кружками для $x > a$.

Уравнение (11) имеет особую точку $x = a$, при приближении к которой рост времени t приводит к малым изменениям переменной x. Поэтому в нашем случае, когда шаг d_x задан постоянным и вычисляется шаг по времени τ, при приближении к особой точке происходит интенсивный рост величины τ, что приводит к повышению погрешности решения. Это можно увидеть на обеих кривых порядка относительной погрешности решения (рис. 2-б).

Полученные результаты вычислений показывают рациональность использования предложенного метода для поиска решения в области интенсивного роста значений функции правой части уравнения. Поэтому разумно предположить эффективность использования этого метода для решения жестких уравнений.

Главным преимуществом представленного метода является его параллельность — на каждом i-м интервале вычисление шага τ_i можно проводить параллельно, после чего решение можно составить $x(t + \sum_{i=1}^{n} \tau_i) = x_0 + nd_x$. А также для нахождения одного шага требуется только одно новое вычисление функции правой части, что по вычислительным затратам сравнимо с методом Эйлера, но точность предлагаемого метода значительно выше.

5. Идея об определении траектории решения автономной системы ОДУ

Рассмотрим автономную систему обыкновенных дифференциальных уравнений первого порядка размерности 2:

$$\begin{cases} \frac{dx_1}{dt} = f(x_1, x_2), \\ \frac{dx_2}{dt} = g(x_1, x_2), \\ x_1(t_0) = x_1^0, \\ x_2(t_0) = x_2^0. \end{cases} \tag{12}$$

Применяя метод разделения переменных к каждому уравнению в отдельности, получим
$$\int \frac{dx_1}{f(x_1,x_2)} = \int dt \Rightarrow t_1(x_1,x_2) = \int \frac{dx_1}{f(x_1,x_2)} + C_1,$$
из второго уравнения
$$t_2(x_1,x_2) = \int \frac{dx_2}{g(x_1,x_2)} + C_2.$$

Введенные здесь функции $t_1(x_1,x_2)$ и $t_2(x_1,x_2)$ описывают поведение интегралов от обратных функций правых частей системы (12). Константы C_1, C_2 определяются, исходя из начальных условий системы.

Предполагаем, что решение системы (12) можно определить как параметрическую кривую Γ в пространстве $\mathbb{R}^2$ вида
$$\Gamma = \{(x_1,x_2) \in \mathbb{R}^2 |\ t_1(x_1,x_2) - t_2(x_1,x_2) = 0\}.$$

Параметр s кривой Γ определяется $s = t_1(x_1,x_2) = t_2(x_1,x_2),\ (x_1,x_2) \in \Gamma$.

Для нахождения численного решения системы (12) необходимо разбить фазовое пространство системы на квадратные участки $S_{ij}, (i \in \mathbb{Z}, j \in \mathbb{Z})$ со стороной d_x:
$$S_{ij} = \{(x_1,x_2) \in \mathbb{R}^2 |\ x_1^0 + id_x \le x_1 \le x_1^0 + (i+1)d_x, x_2^0 + jd_x \le x_2 \le x_2^0 + (j+1)d_x\}.$$

Для каждого квадрата S_{ij} вычисляются значения функций правых частей системы в углах этого квадрата. Используя билинейную интерполяцию, вычисляются интегралы для определения t_1, t_2. Зная точку входа траектории решения в квадрат, определяются константы C_1 и C_2, после чего находится точка выхода траектории решения. Эта точка, в свою очередь, является точкой входа в соседний квадрат, где процедура повторяется. Таким образом, траектория решения определяется точками, на границах квадратов.

Отметим, что вычисление интегралов в каждом квадрате S_{ij} является независимым от значений функций в других квадратах, что дает возможность производить вычисления в параллельном режиме. Определение конкретной траектории решения, соответствующей заданной начальной точке, будет основано на определении точек выхода для каждой клетки, для чего потребуется находить только константы C_1 и C_2 по заранее вычисленным интегралам. Причем эти интегралы не нужно вычислять повторно, если необходимо найти другую траекторию, достаточно вычислить константы. Что дает ускорение в определении фазового портрета системы.

Для увеличения точности определения точки пересечения границы необходимо заменить билинейную интерполяцию на биквадратичную, бикубическую и т.д.

Предложенную идею определения траектории решения систему ОДУ можно распространить на n-мерный случай. Траекторий будет определяться как пересечение $n-1$ гиперповерхностей размерности $n-1$.

Литература

1. Деккер, К. Устойчивость методов Рунге-Кутты для жестких нелинейных дифференциальных уравнений / К. Деккер, Я. Вервер. – М.: Мир, 1988.
2. Коробицын, В.В. Алгоритм численного решения дифференциальных уравнений с разрывной правой частью / В.В. Коробицын, Ю.В. Фролова // Математические структуры и моделирование. – 2005. – Вып. 15. – С.46–54.
3. Коробицын, В.В. Исследование поведения явных методов Рунге-Кутты при решении систем обыкновенных дифференциальных уравнений с разрывной правой частью / В.В. Коробицын, В.Б. Маренич, Ю.В. Фролова // Математические структуры и моделирование. – 2007. – Вып. 17. – С.19–25.
4. Коробицын, В.В., Алгоритм численного решения кусочно-сшитых систем / В.В. Коробицын, Ю.В. Фролова, В.Б. Маренич // Вычисл. технологии. – 2008. – Т.13, N.2. – С.70–81.
5. Новиков, Е.А. Реализация полуявного (4,2)-метода решения жестких систем на параллельной ЭВМ / Е.А. Новиков, Л.П. Каменщиков // Вычисл. технологии. – 2004. – Т.9. Вестник КазНУ им. аль-Фараби. Сер. Математика, механика, информатика. N.3 (42). (Совм. выпуск. Ч.1). – С.235–241.
6. Фельдман, Л.П. Сходимость и оценка погрешности параллельных одношаговых блочных методов моделирования динамических систем с сосредоточенными параметрами / Л.П. Фельдман // Научные труды ДонНТУ. Серия: Информатика, моделирование и вычислительные методы, (ИКВТ-2000), выпуск 15. – Донецк: ДонНТУ, 2000. – С.34–39.
7. Фельдман, Л.П. Параллельные алгоритмы численного решения задачи Коши для систем обыкновенных дифференциальных уравнений / Л.П. Фельдман, И.А. Назарова // Матем. моделирование. – 2006. – Т.18, N.9. – С.17–31.
8. Хайрер, Э. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи / Э. Хайрер, С. Нёрсетт, Г. Ваннер. – М.: Мир, 1990.
9. Хайрер, Э. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи / Э. Хайрер, Г. Ваннер. – М.: Мир, 1999.
10. Cash, J.R. A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides / J.R. Cash, A.H. Karp // ACM Trans. Math. Software. – 1990. – V.16, N.3. – P.201–222.
11. Gear, C.W. Solving ordinary differential equations with discontinuities / C.W. Gear, O. Østerby // ACM Trans. Math. Software. – 1984. – V.10. P.23–44.
12. Jackson, K. The potential for parallelism in Runge-Kutta methods. Part I: RK formulas in standard form / K. Jackson, S. Norsett // SIAM J. Numer. Anal. – 1995. – V.32. – P.49–82.

Математические
структуры и моделирование
2009, вып. 19, с. 28–44

УДК 000.000

ЭФФЕКТИВНЫЙ ТОПОС КАК СИНТЕТИЧЕСКАЯ ВСЕЛЕННАЯ ДЛЯ ТЕОРИИ ВЫЧИСЛЕНИЙ

Е.О. Хлызов

В статье обсуждается применение теории элементарных топосов Ловера к теории вычислимости. Дается пример построения синтетической теории вычислений на основе так называемого эффективного топоса.

Краткое вступление

Предположим, некий математик собрался изучить вполне определенный вид математических структур, таких как, например, топологические пространства в паре с непрерывными отображениями, или эффективно построимые множества в совокупности с эффективно вычислимыми функциями. Для выделения указанных объектов, математик следует классическому пути, рассматривая множества, наделенные дополнительной структурой и отображения, сохраняющие эту структурур. Тогда можно говорить о том, что он изучает топологию и теорию вычислений. Но есть и другой путь – изолировать изучаемые объекты, поместив их в некоторую синтетическую вселенную, в которой специфическая структура объекта является всеобщей, т.е. распространяющейся на всю вселенную. Таким образом, мы получаем синтетический мир с некоторой заданной в нем логикой (не обязательно классической). Находящийся в этом мире математик, со своей точки зрения, изучает простые множества и простые отображения, а с точки зрения внешнего наблюдателя изучает строго определенный класс множеств и отображения, позволяющие остаться в рамках этого класса. Такой подход принято называть синтетическим.

Автор рассматривает данную статью, как попытку в максимально простой форме, на примере теории вычислимости, показать возможности использования топосов в качестве синтетических вселенных. В ней рассматривается пример построения синтетической теории вычислений на основе так называемого *эффективного топоса*.

Омский государственный университет им. Ф.М. Достоевского.
E-mail: `hlyzov@gmail.com`

1. Используемые обозначения

Функцию, отображающую заданное множество на множество всех его подмножеств, будем обозначать $\mathcal{P}$. Под A^B подразумевается множество всех функций из B в A. Обозначение ! для стрелки отражает факт ее единственности. Также, для выражения «$S1$ такое что $S2$» будет использоваться сокращение «$S1.S2$». Истина и ложь обозначаются знаками $\perp$ ($false$) и $\top$ ($true$) соответственно. Объяснения остальных обозначений будут даны по ходу изложения материала.

2. Начальные определения

Для понимания следующих параграфов от читателя потребуется знание основных определений теории категорий. Далее приводится необходимый минимум, для того, чтобы подойти к понятию элементарного топоса. Для более детального ознакомления с теорией категорий и теорией топосов рекомендуется обратится к соответствующей литературе (см. например [1–3]).

Определение 1. Категория $\mathcal{C}$ включает в себя:

1. Коллекцию **объектов** $Ob(\mathcal{C})$.

2. Коллекцию **морфизмов** (*стрелок*) $Mor(\mathcal{C})$.

3. Две операции $dom, cod : Mor(\mathcal{C}) \to Ob(\mathcal{C})$ сопоставляющие каждому морфизму f два объекта, называемые, соответственно, **областью определения** (*доменом*) и **областью значений** (*кодоменом*) морфизма f.

4. Оператор композиции $\circ : Mor(\mathcal{C}) \times Mor(\mathcal{C}) \to Mor(\mathcal{C})$ сопоставляющий каждой паре морфизмов $\langle f, g\rangle$, таких что $dom(g) = cod(f)$ морфизм $g \circ f : dom(f) \to cod(g)$, таким образом, что выполняется **закон ассоциативности**:

$$h \circ (f \circ g) = (h \circ f) \circ g$$

5. Оператор $id : Ob(\mathcal{C}) \to Mor(\mathcal{C})$, опоставляющий каждому объекту A, тождественный морфизм $id_A : A \to A$, таким образом что выполняется **закон единицы**:

$$id_B \circ f = f = f \circ id_A$$

Примеры категорий:

1. **Set** – объекты: множества, стрелки: все функции между множествами.

2. **Top** – объекты: топологические пространства, стрелки: все непрерывные функции.

3. **N** – объект: N (единственный), стрелки: отображения из N в N. Каждая стрелка представляет собой натуральное число ($c = m \circ n = m + n \in Mor(\mathbf{N}), \forall m, n \in Mor(\mathbf{N})$).

Определение 2 (Декартово замкнутая категория)**.** Категория $\mathcal{C}$ декартово замкнута, если она:

1. Содержит терминальный объект, т.е. объект к которому каждый объект категории имеет единственную стрелку, иначе говря, утверждается, что: $\exists\ \mathbf{1} \in Ob(\mathcal{C})\ .\ \forall X \in Ob(\mathcal{C})\ \exists! f \in Mor(\mathcal{C})\ .\ cod(f) = \mathbf{1}, dom(f) = X$. В категории Set примером терминального объекта является множество с одним элементом.

2. Для любых двух объектов $X, Y \in Ob(\mathcal{C})$ существует объект $X \times Y \in Ob(\mathcal{C})$ который, вместе с двумя стрелками $pr_X : X \times Y \to X$, $pr_Y : X \times Y \to Y$, обладает следующим свойством: для произвольной пары стрелок $f : C \to X$, $g : C \to Y$ существует единственная стрелка $h : C \to X \times Y$, такая что следующая диаграмма коммутативна:

$$\begin{tikzcd} & C \arrow[dl, "f"'] \arrow[d, dashed, "h"'] \arrow[dr, "g"] & \\ X & X \times Y \arrow[l, "pr_X"'] \arrow[r, "pr_Y"] & Y \end{tikzcd}$$

   Тройку $< X \times Y, pr_X, pr_Y >$ называют **категорным произведением** X и Y, а стрелки pr_X и pr_Y называют **проекциями**. В категории Set категорным произведением является прямое произведение двух множеств.

3. Категория допускает **экспоненцирование**, т.е. для любых двух объектов $X, Y \in Ob(\mathcal{C})$ существует объект $Y^X \in Ob(\mathcal{C})$ называемый экспоненциалом и специальная стрелка $ev : Y^X \times X \to Y$ называемая **функцией значения**, такие, что для любого объекта $Z \in Ob(\mathcal{C})$ и любой стрелки $g : Z \times X \to Y$ существует единственная стрелка $\hat{g} : Z \to Y^X$, такая что следующая диаграмма коммутативна:

$$\begin{tikzcd} Y^X \times X \arrow[dr, "ev"] & \\ Z \times X \arrow[u, dashed, "\hat{g} \times 1_X"] \arrow[r, "g"'] & Y \end{tikzcd}$$

   В категории Set экспоненциалом Y^X является множество всех функций из X в Y, а функцией значения – отображение, переводящее пару ($f \in Y^X, X$) в множество $f(X)$.

Определение 3 (Классификатор подобъектов)**.** Говорят что категория $\mathcal{C}$ с терминальным объектом **1** имеет **классификатор подобъектов** Ω, если существует единственная стрелка $\top : \mathbf{1} \to \Omega$ и выполняется следующее правило (Ω-аксиома):

Для каждого инъективного отображения (моностредки) $f : A \to B$ существует единственная $\chi_A : B \to \Omega$ такая что: $\top \circ ! = \chi_f \circ f$. На языке теории категорий предыдущее выражение формулируется в виде утверждения, что следующая диаграмма **коммутативна** (коммутативные квадраты принято называть *декартовыми*):

$$\begin{tikzcd} a \arrow[r, "f"] \arrow[d, "!"'] & d \arrow[d, "\chi_A"] \\ \mathbf{1} \arrow[r, "\top"'] & \Omega \end{tikzcd}$$

Стрелку χ_A при этом называют **характеристической** стрелкой для A. В категории Set, Ω имеет вид $\{0,1\}$, а характеристическая функция для подмножества $A \subset B$ определяется следующим образом:

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Наконец, мы подошли к наиболее интересующему нас понятию топоса. Следующее определение эквивалентно данному В. Ловером и М. Тирни в 1969 году:

Определение 4 (Топос)**.** Декартово замкнутая категория с классификатором подобъектов называется ***элементарным топосом***.

Слово «элементарный» лишь указывает на природу данного топоса, а точнее на ее отличие от известного до него **топоса Гротендика** и в дальнейшем будет опускаться.

К сожалению, без полноценного введения в теорию категорий сложно привести примеры топосов отличных от Set. Поэтому ограничимся простейшими топосами, так или иначе образованными от Set:

1. Самым естественным примером топоса является сама категория Set (вообще, топос можно мыслить как некоторое обобщение Set). В определениях выше именно эта категория была использована в качестве наглядного пособия.

2. Категория Finset является подкатегорией категории Set и в качестве объектов содержит конечные множества. Finset также является топосом – экспоненциал, классификатор и категорное произведение строятся также как в Set.

3. Категория пучков множеств в топологическом пространстве является топосом [5]. Вообще, категория множеств является топосом, как категория пучков множеств одноточеченого пространства.

Более сложный пример топоса, позволяющего работать с «эффективно конструируемыми» множествами и «эффективно вычислимым» отображениями, будет рассмотрен чуть позже. Топосы интересеы тем, что они могут выступать в качестве описанной в начале статьи синтетической вселенной. Впервые такие эксперименты были проведены Ловером, что привело к созданию синтетической дифференциальной геометрии [8, 9].

3. Зарождение синтетической теории вычислимости

В связи с тем, что каждый топос определяет собственное логическое исчисление [2], упоминание о синтетической теории вычислимости (СТВ) можно встретить еще в 1986 г. [11], т.е. практически одновременно с появлентем **эффективного** и **рекурсивного** топосов. В то время СТВ не получила значимого развития и «категорификация» приближалась к теории вычислений посредством *синтетической теории доменов* – теории, где применение синтетического подхода позволило получить многообещающие результаты [**?**, 12]. В 2005 году словенский математик Андрей Бауер выступил с докладом «Первые шаги в синтетической теории вычислимости», где была сделана попытка определить начальные положения новой теории, опираясь на уже довольно богатый опыт наработок с вычислениями в категориях. В качестве модели он предпочел использовать эффективный топос *Eff*, описанный М. Хиландом в [7]. Интересно, что *Eff* является элементарным топосом, но не является топосом Гротендика, в отличии от своего «конкурента» – рекурсивного топоса [11]. Далее мы »вскроем» СТВ предложенную Бауером и разберем строение *Eff*.

4. Строение категории *Eff*

Рассмотрим новую категорию «предпорядок». Определенную тем, что для любых двух ее объектов $X, Y \in P$ существует не более одного морфизма $f : X \to Y$. На множестве объектов P (в случае когда можно говорить о совокупности объектов именно как о множестве, т.е. категория P **мала**) можно задать отношение предпорядка следующим образом: $X \leq Y \iff \exists f : X \to Y$.

Следующее определение которое понадобится далее относится к категорным построениям и получается путем обращения всех стрелок в определении категорного произведения.

Определение 5 (Копроизведение)**.** Копроизведением двух объектов $X, Y \in Ob(\mathcal{C})$ называется объект $X+Y \in Ob(\mathcal{C})$ который, вместе с двумя стрелками $i_X : X \to X+Y$, $i_Y : Y \to X+Y$, обладает следующим свойством: для произвольной пары стрелок $f : X \to C$, $g : Y \to C$ существует единственная стрелка h :

$X+Y \to C$, такая что следующая диаграмма коммутативна:

$$\begin{tikzcd} & C & \\ X \arrow[ur, "f"] \arrow[r, "i_X"] & X+Y \arrow[u, dashed, "h"] & Y \arrow[ul, "g"'] \arrow[l, "i_Y"'] \end{tikzcd}$$

Тройку $< X+Y, i_X, i_Y >$называют **категорным копроизведением** или **суммой** X и Y, а стрелки i_X и i_Y называют **инъекциями**. В категории Set категорным произведением является прямое произведение двух множеств.

Для построения *Eff* нужна алгебра более общая чем алгебра Гейтинга (которая определяется на частично упорядоченных множествах), но имеющая схожие свойства. Заменой отношение частичного порядка на отношение предпорятка, получаем следующее определение.

Определение 6. Предалгебра Гейтинга H – это декартово замкнутая категория предпорядка с произведениями ($\wedge$), копроизведениями ($\vee$), начальным объектом $\perp$, конечным объектом $\top$ и экспоненциалом ($\to$). Под $p \equiv q$ подразумевается, что $p \leq q$ и $q \leq p$.

Легко проверить, что предалгебра Гейтинга является интерпретацией конструктивной логики высказываний [10]. Конструктивность можно мыслить как необходимость «подтверждать» абстрактные выводы, т.е. для каждого такого вывода должен предъявляться эффективный алгоритм вычисления результата. В конструктивной логике высказывание $(A \vee !A)$ не доказуемо, чтобы понять почему, достаточно представить себе множество «точек останова» $K \subset \mathbb{N}$. Если конструктивной истинной было бы высказывание $n \in K \vee n \notin K$ то это означало бы, что существует алгоритм, который, будучи примененным на натуральное число n, сообщил бы, входит ли n в K или нет. Но, согласно тезису Черча, такого алгоритма не существует.

Определение 7. Пусть X – некоторое множество, тогда мы можем превратить степенное множество $\mathcal{P}(\mathbb{N})$ в предалгебру Гейтинга следующим способом:

Если $F, G \in (\mathcal{P}(\mathbb{N}))^X$, таким образом что $F(x), G(x) \subseteq \mathbb{N}$ для кажого $x \in X$, то:

$$\begin{gathered} (F \wedge G)(x) = F(x) \wedge G(x) \\ (F \vee G)(x) = F(x) \vee G(x) \\ (F \to G)(x) = F(x) \to G(x) \end{gathered}$$

И $F \leq G \iff \bigcap\limits_{x \in X} F(x) \to G(x) \neq \varnothing$

Также мы будем писать $\mathcal{P}(\mathbb{N}) \vDash (F \leftrightarrow G)$, для обозначения $\mathcal{P}(\mathbb{N}) \vDash (F \to G) \wedge (G \to F)$.

Определение 8. Пусть $F \in (\mathcal{P}(\mathbb{N}))^X$, если выполняется $\bigcap_{x \in X} F(x) \neq \varnothing$, то будем это записывать как $\mathcal{P}(\mathbb{N}) \vDash F$, а выполнение $F = \lambda x.F(x)$ будем записывать как $\mathcal{P}(\mathbb{N}) \vDash F(x)$.

Теперь можно перейти непосредственно к описанию элементов категории *Eff*. Доказательство единственной в этом параграфе теоремы (а также многих других, не вошедших в данную статью) можно посмотреть в соответствующей монографии [7].

Определение 9 (Объект *Eff*)**. Объект** $(X, =)$ в *Eff* – это множество X вместе с отношением эквивалентности в $\mathcal{P}(\mathbb{N})$, т.е. для каждых $x, x' \in X$, подмножество $|x = x'| \subseteq \mathbf{N}$ реализуемо симметрично и транзитивно, что означает:

$$\mathcal{P}(\mathbb{N}) \vDash |x = x'| \to |x' = x| \text{ и}$$
$$\mathcal{P}(\mathbb{N}) \vDash (|x = x'| \wedge |x' = x''|) \to |x = x''|$$

Пусть $x \in X$, будем писать Ex для обозначения $|\mathrm{x} = \mathrm{x}|$.

Определение 10 (Функциональное отношение в *Eff*)**.** Пусть $(X, =), (Y, =) \in Ob(\mathit{Eff})$, тогда функциональное отношение из $(X, =)$ в $(Y, =)$ – это отображение $F : X \times Y \to \mathcal{P}(\mathbb{N})$, такое что

- $\mathcal{P}(\mathbb{N}) \vDash (F(x, y) \wedge (|x = x'| \wedge |y = y'|)) \to F(x', y')$
- $\mathcal{P}(\mathbb{N}) \vDash F(x, y) \to (Ex \wedge Ey)$
- $\mathcal{P}(\mathbb{N}) \vDash F(x, y) \wedge F(x, y') \to |y = y'|$
- $\mathcal{P}(\mathbb{N}) \vDash Ex \to \{\langle k, m\rangle \,|k \in Ey \;\; y \in Y \;\; m \in F(x, y)\}$

два функциональных отношения F, F' **эквивалентны** $(F \sim F')$, если

$$\mathcal{P}(\mathbb{N}) \vDash (F(x, y) \longleftrightarrow F'(x, y))$$

Определение 11 (Тождественность в *Eff*)**.** Тождественное функциональное отношение в *Eff* из $(X, =)$ в $(X, =)$ определяется как $I(x, x') = |x = x'|$

Определение 12 (Морфизм *Eff*)**. Морфизм**(стрелка) $f : (X, =) \to (Y, =)$ в *Eff* – это класс эквивалентности функциональных отношений $f = [F]$, где F – функциональное отношение, задающее f.

Теорема 1 (Хайланд)**.** *Eff – топос.*

Для лучшего уяснения внутреннего строения эффективного топоса рекомендуется обратится к соответствующей литературе [7, 10, 11]. Далее, чтобы избежать серьезного углубления в теорию топосов (беря пример с Бауера) будет использоваться специальный внутрений язык. Внешнее описание *Eff* было дано выше, внутренне же топос можно представить как конструктивную теорию множеств, расширенную несколькими аксиомами, касающимися множеств и последовательностей натуральных чисел [4, 6].

5. Введение в Синтетическую теорию вычислений

В качестве одной из причины возникновения СТВ можно указать некоторую обособленность теории вычислений от остальной математики. Описания различных машин Тьюринга, числа Геделя, лямбда-абстракции придают ей некоторую особенность, но в то же время поднимают порог вхождения, требуя от математика знания результатов полученных при использовании той или иной модели вычислений. СТВ позволяет работать с гораздо более привычными объектами – множествами и отображениями между ними, скрывая детали за высоким уровнем абстракции с помощью «внутреннего» языка родительского топоса. Поскольку в данном случае в качестве фреймворка используется топос *Eff* и одна из аксиом СТВ противоречит законы исключенного третьего в классической логике, мы будем работать в интуиционистской (конструктивной) логике. Наша цель на данном этапе – перенести некоторые понятия и теоремы классической теории на, на внутренний язык *Eff* не выходя далеко за пределы конструктивной теории множеств. Также следует отметить, что в *Eff* вычислимость является встроенной, т.е. все что выводимо в *Eff* по определению имеет эффективный алгоритм для вычисления. Дальнейший текст является в большей степени обзорным, предложения и теоремы приводятся без доказательства. Для более полного ознакомления следует обратится к [4].

То, что мы работаем в интуиционистской логике (ИЛ), означает, что мы будем иметь дело с выражениями выходящими за рамки «классического» понимания. Помимо закона исключенного третьего, следующие общепринятые формулы также не верны (в общем) в ИЛ:

1. $\neg\neg\varphi \Longrightarrow \varphi$
2. $\neg\varphi \vee \neg\neg\varphi$
3. $(\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi)$
4. $\neg(\neg\varphi \wedge \neg\psi) \Longrightarrow \varphi \vee \psi$

В частности, это означает, что в интуиционистском смысле существуют аргументы на которых ни одна из выше приведенных формул не выполняется. Отдельного параграфа заслуживает аксиома выбора.

6. Строение множеств

Для определения упорядоченной пары (элемент множества $A \times B$) мы рассмотрим операцию $\langle -, - \rangle$ вместе с проекциями π_1 и π_2, удовлетворяющих акиомам: $\pi_1 \langle x, y \rangle = x$ и $\pi_2 \langle x, y \rangle = y$. Определим понятия суммы и произведения семейства множеств $\mathcal{F} \subseteq \mathcal{P}(A)$:

$$\cup\mathcal{F} = \{x \in A | \exists S \in \mathcal{F}. x \in S\} \cap \mathcal{F} = \{x \in A | \forall S \in \mathcal{F}. x \in S\}$$

Теперь мы готовы для следующего важного определения:

Определение 13 (Натуральные числа)**.** Множество $\mathbb{N}$ вместе с элементом $0 \in \mathbb{N}$ называющимся *нулем* и с функцией *выбора следующего элемента* $succ : \mathbb{N} \to \mathbb{N}$, такой что для каждого множества A, $x \in A$ и $f : A \to A$ существует единственная функция $h : \mathbb{N} \to A$, такая что $h(0) = x$ и $h(succ(n)) = f(h(n)$. Будем говорить что функция h задана **простой рекурсией**. Одноэлементное множество $\mathbf{1}$ определяется как $\mathbf{1} = \{n \in \mathbb{N} | n = 0\}$. А пустое множество $\emptyset$ как $\emptyset = \{x \in \mathbf{1} | \bot\}$.

7. Строения функций

Граф $\Gamma(f)$ функции $f : A \to B$ – это отношение $\Gamma(f) \subseteq A \times B$ определенное для $x \in A, y \in B$ следующим образом:

$$(x, y) \in \Gamma(f) \Longleftrightarrow f(x) = y$$

Легко проверить что граф является функциональным отношением. Обратно, каждое функциональное отношение определяет функцию. Это известно как *аксиома единственного выбора*:

$$\forall R \in A \times B.((\forall x \in A.\exists! y \in B.R(x, y)) \Longrightarrow \exists f \in B^A.\forall x \in A.R(x, f(x)))$$

Приведенная аксиома оказывается полезной при задании функций. Например единичная функция $\mathbf{1}_A$ определяется как функция, чей граф является отношением равенства на A.

Теперь имеется достаточно инструментов для того, чтобы приступить к описанию теории. Прежде чем перейти к рассмотрению основных аксиом СТВ, определим понятие *разрешимости.* Для начала посмотрим на понятие истинностного значения в контексте ИЛ. Напомним, что истинностным значением называется высказывание без свободных переменных (например, $\top$, $\bot$, $\forall x \in (R).(x = 0 \vee x \neq 0)$).

Определение 14 (Множество истинностных значений)**.** $\Omega = P\mathbf{1}$

В классической логике показывается, что только $\bot$ (ложь) и $\top$ (истина) являются элементами Ω (в некотором роде это просто форма закона исключенного третьего). В интуиционсткой логике это утверждение неверно. Впрочем, это **не** означает что Ω состоит более чем из двух элементов... Для того, чтобы это понять, следует вспомнить что означает высказывание «множество имеет два элемента»: если $x, y \in A$, такие что $x \neq y$ и не существует элемента $z \in A$, такого что $z \neq x$ и $z \neq y$ – говорят что множество A имеет два элемента в *слабом смысле.* С другой стороны, если при заданных условиях для любого элемента $z \in A$ верно, что либо $z = x$, либо $z = x$, говорят что множество A имеет два элемента в сильном смысле. Таким образом Ω имеет два элемента ($\bot$ и $\top$) в слабом смысле, но доказать, что Ω имеет два элемента в строгом смысле мы можем только в рамках классической логики.

Ω является предалгеброй Гейтинга [1]. Таким образом $p \wedge q \Longrightarrow\Longleftrightarrow p \Longrightarrow$

[1]На самом деле, как показательное множество, Ω является полной алгеброй Гейтинга

$(q \Longrightarrow r)$. Отрицание является псевдо-дополнением, т.е. $\neg p$ определяется как наибольший элемент q такой что $p \wedge q = \ \perp$. Рассмотрим подмножество Ω, имеющее отношение к классической логике, элементы которого *разрешимы*, т.е. удовлетворяют закону исключенного третьего.

Определение 15 (Множество разрешимых истинностных значений)**.**

$$\mathbf{2} = \{p \in \Omega | p \vee \neg p\}$$

2 имеет ровно два элемента в строгом смысле и вместе с операциями $\wedge$ и $\vee$ является булевой алгеброй.

Представим еще одно подмножество Ω имеющее непосредственное отношение к классической логике

Определение 16 (Классическое множество истинностных значений)**.**

$$\Omega_{\neg\neg} = \{p \in \Omega | \neg\neg p \Longrightarrow p\}$$

Элементами $\Omega_{\neg\neg}$ являются истинностные значения, чья истинность может быть доказана «от противного» – если $\neg p$ ложно значит p истинно. Дальше в тексте элементы $\Omega_{\neg\neg}$ будут называться «классическими». $\Omega_{\neg\neg}$ помимо того, что как и Ω является полной алгеброй Гейтинга, также является полной булевой алгеброй.

Приведенные множества соотносятся следующим образом.

$$\mathbf{2} \subseteq \Omega_{\neg\neg} \subseteq \Omega$$

Более полную информацию можно будет получить позднее, когда добавится третье подмножество $\Sigma \subseteq \Omega$, играющее ключевую роль в СТВ.

8. Предикаты

Предикат $P \subseteq A$ может быть задан характеристической функцией $\chi_P : A \to \Omega$, определенной следующим образом:

$$\chi_P(x) = \{t \in \mathbf{1} | x \in P\}$$

Обратно, функция $\xi : A \to \Omega$ соотносится с подмножеством

$$\{x \in A | \ \xi(x) = \top\} \in A$$

Таким образом установлено биективное соответствие между предикатами на , подмножествами и пропозициональными функциями $A \to \Omega$. Пропозициональные функции отображающие на подмножества Ω приводят к особым видам предикатов. Например функция $p : A \to \mathbf{2}$ представляет подмножество $S = \{x \in A | \ p(x) = \top\}$, такое что $\forall x \in A.(x \in S \vee x \notin S)$ Такие предикаты и подмножества называются *разрешимыми*.

Если равенство на множестве задано вычислимым предикатом $(\forall x \in A.(x = y \vee x \neq y))$, то такое множество называется вычислимым.

Предложение 1. *Следующие множества разрешимы: натуральные числа, подмножества разрешимого множества, декартово произведение и сумма разрешимых множеств.*

9. Аксиомы СТВ

Определение 17 (Проективное множество)**.** Говорят, что два множества A, B *допускают выбор* (обозн. $AC(AB)$), если для каждого тотального отношения между A и B выполняется

$$\forall R \subseteq A \times B.((\forall x \in A.\exists y \in B.R(x,y)) \Longrightarrow \exists f \in B^A.\forall x \in A.R(x, f(x)))$$

Множество называется проективным, если $AC(A, B)$ выполняется для каждого множества B.

В классической теории множеств акисома выбора утверждает, что все множества проективны. В нашем случае мы ограничены гораздо более серьзно, потому что в вычислительном смысле, проективными являются только множества, каждый элемент которых имеет канонический код Геделя. Таким образом, можно ожидать, что множество $\mathbb{N}$ проективно, а множество $\mathbb{N}^{\mathbb{N}}$ – нет, потому что в последнем случае нельзя эффективным образом указать процедуру выбора канонического кода.

В конструктивной математике множество натуральных числе несомненно должно быть проективно. Это уставнавливает следующая аксиома.

Аксиома 1 (Выбора числа)**.** *Множество натуральных чисел* $\mathbb{N}$ *проективно.*

Также понадобится следующее обобщение аксиомы выбора числа.

Аксиома 2 (Обусловленный выбор)**.** *Если R тотальное отношениа на A и $x \in A$, тогда существует $f : \mathbb{N} \to A$, такое что $f(0) = x$ и $R(f(n), f(n+1))$ выполняется для всех $n \in \mathbb{N}$.*

Третья аксиома выбора одновременно является и первой аксиомой СТВ. В конструктивной математике (по Бишопу) она в общем не принимается. Ее обоснование лежит на рассмотрении вычислительного аспекта классических подмножеств. Членство в классическом подмножестве $S \subseteq A$ проективного множества не несет никакой вычислительной нагрузки, таким образом коды Геделя для элементов A кодируют ту же информациюю, что и для элементов S. Таким образом, мы можем использовать те же коды для множества и его подмножества, а раз элементы A имеют канонические кода, то значит элементы S также будут их иметь.

Аксиома 3 (Проективности)**.** *Классическое подмножество проективного множества проективно.*

10. Счетные и полусчетные множества

Вычислимые счетные множества играют центральную роль в классической теории вычислений. В СТВ они сохраняют свою значимость, только внутри *Eff* они выглядят как обычные счетные множества.

Определение 18. Множество A счетно (перечислимо), если существует сюрьективное отображение $e : \mathbb{N} \to \mathbf{1} + A$, называемое нумерацией A. Нумерация называется нумерацией без повторений, если $\forall n, m \in \mathbb{N}, e(n) = e(m) \Rightarrow n = m$.

Предложение 2. *Множество может быть пронумеровано без повторений тогда и только тогда когда она счетно и разрешимо.*

Следствие 1.1. *Каждое счетное подмножество* (N) *может быть пронумеровано без повторений.*

Ниже представлены некоторые утверждения касательно способов создания счетных множеств.

1. Декартово произведение двух счетных множеств счетно.
2. Объединение счетного семейства множеств счетно.
3. Пересечение двух счетных подмножеств разрешимого множества счетно.
4. Разрешимое подмножество счетного подмножества счетно.
5. Конечная последовательность счетных множеств формирует счетное множество.

Определенный интерес представляет собой множество всех счетных подмножеств $\mathbb{N}$, будем обозначать его $\mathcal{E}$:

$$\mathcal{E} = \{A = \mathcal{P}\mathbb{N} | A\ \}.$$

С этим множеством связана следующая любопытная теорема, показывающая что $\mathcal{E}$ является дискретной топологией на $\mathbf{N}$

Теорема 2. *Семейство* (E) *это наименьшее из семейств* $(F) \subseteq \mathcal{P}(\mathbb{N})$ *таких что:*

1. $0 \in (F)$ *и* $\mathbf{N} \in \mathcal{F}$.
2. $\{n\} \in \mathcal{F}$*для каждого* $n \in \mathbf{N}$.
3. *Если* $A, B \in \mathcal{F}$ *то* $A \cap B \in \mathcal{F}$.
4. *Если* $\mathcal{A} \in \mathcal{F}$*счетное семейство, тогда* $\cup\mathcal{A} \in \mathcal{F}$.

Теорема 3 (Теорема о проекции)**.** *Помножество* $S \subseteq \mathbb{N}$ *счетно тогда и только тогда, когда оно является проекцией разрешимого подмножества декартового произведения* $\mathbb{N} \times \mathbb{N}$ *. Т.е.* $S = \{x \in \mathbb{N} | \exists y \in \mathbb{N}. \langle x, y\rangle \in \mathbb{N} \times \mathbb{N}\}$

Следующее подмножество занимает исключительную роль в СТВ, а также хорошо известно в синтетической теории доменов. Мы покажем, что в СТВ счетные множества являются *полуразрешимыми*, предоставив такое множество $\Sigma \subseteq \Omega$ полуразрешимых истинных значений, что $\mathcal{E} = \Sigma^{\mathbb{N}}$.

Определение 19 (Множество полуразрешимых истинностных значений)**.**

$$\Sigma = \{p \in \Sigma | \exists f \in \mathbf{2}^{\mathbb{N}}.(p \iff (\exists n \in \mathbb{N}.f(n)))\}.$$

Σ также является *доминирующим* (см. [11]):

$$\forall p \in \Sigma.\forall q \in \Omega.((p \Rightarrow (q \in \Sigma)) \implies (p \wedge q) \in \Sigma.$$

Приведем несколько утверждений, раскрывающих значение Σ.

Предложение 3. *Подмножество* $\mathbb{N}$ *счетно тогда и только тогда, когда она полуразрешимо.*

Предложение 4. Σ *это наименьшее подмножество* Ω *содержащее* $\top$ *и замкнутое сверху.*

То, как эти множества соотносятся между собой, хорошо показывает следующая теорема (для доказательства см. [4])

Теорема 4. *Ни одно из следующих включений* $\mathit{2} \subseteq \Sigma \subseteq \Omega_{\neg\neg} \subseteq \Omega$ *не является равенством.*

В классическое логике все несколько проще: $\mathit{2} = \Omega$ со всеми вытекающими из этого равенства последствиями.

Рассмотрим широко известный в классической теории вычислений принцип Маркова.

Предложение 5. *Следующий утверждения эквивалентны:*

1. *Принцип Маркова: для каждого* $a : \mathbb{N} \to \mathbf{2}, \neg(\forall n \in \mathbb{N}.a_n = 0) \implies \exists n \in \mathbb{N}.a_n = 1$
2. *Полуразрешимые истинностные значения являются классическими* $\Sigma \subseteq \Omega_{\neg\neg}$
3. *Полуразрешимые подмножества являются классиескими.*
4. *Полуразрешимые подмножества* $\mathbb{N}$ *являются классическими.*

Хотя его выполнение является интуитивно бесспорным, доказать это в рамках конструктивной теории невозможно.

Аксиома 4 (Принцип Маркова)**.** *Двоичная последовательность не являющаяся нулем, содержит единицу.*

С помощью этой аксиомы можно доказать следующую теорему, приводимую практически в каждой книге по теории вычислимости (по крайней мере для частного случая $A = \mathbb{N}$)

Теорема 5 (Поста)**.** *Подмножество множества* A *разрешимо тогда и только тогда, когда дополнение этого подмножества полуразрешимо.*

11. Основы СТВ

В этом параграфе будет дана последняя аксиома СТВ и несколько базовых теорем теории вычислений (которые удалось доказать в СТВ).

Определение 20. *Частичная* функция $f : A \rightharpoonup B$ это функция $f : A' \to B$ определенная на подмножестве $A' \subseteq A$, называемом *носителем* f Каждая частичная функция f соответствует (тотальной) функции $g : A \to \tilde{B}$, где $\tilde{B}$ это множество частичных значений

$$\tilde{B} = \{\mathcal{P}(B) | \forall x, y \in B.(x \in s \wedge y \in s \Longrightarrow x = y)\}$$

f и g соотносятся следующим образом: $g(x) = \{f(x) \in B | x \in A'\}$. Одноэлементное множество $\{y\}$ представляет собой полностью определенный элемент $y \in B$, который называется *совершенным* значением. Утверждение $\exists y \in B.(s = \{y\})$, означающие „частичное значение s совершенно“ будем сокращать как $s \downarrow$.

Предложение 6. *Частичная функция $\mathbb{N} \to \tilde{\mathbb{N}}$ счетна (имеет счетный граф) тогда и только тогда, когда $f(n) \downarrow$ полуразрешимо для каждого $n \in \mathbb{N}$.*

Рассмотрим те частичные значения, чья совершенность полуразрешима.

Определение 21. Подъемом $A_\perp$ называется множество Σ-частичных значений

$$A_\perp = \{s \in \tilde{A} | s \downarrow \in \Sigma\}$$

Σ-частичная функция это частичная функция $f : A \to B_\perp$.

Σ-частичные функции $\mathbb{N} \to \mathbb{N}_\perp$ являются синтетическими аналогими частично вычислимых функций. Классическая теорема теории вычислимости утверждает, что вычислимо-счетные множества в точности являются носителями частично вычислимых функций. Приведем аналогичное утверждение:

Предложение 7.

1. *Частиная функция является Σ-частичной тогда и только тогда, когда ее носитель полуразрешим.*

2. *Подмножество полуразрешимо тогда и только тогда, когда она является носителем Σ-частичной функции.*

Представим еще одну хорошо известную теорему из классической теории.

Теорема 6 (Теорема об однозначности)**.** *Каждое полуразрешимое отношение R на $\mathbb{N} \times \mathbb{N}$ имеет Σ-частичную функцию выбора $s : A \to \tilde{B}$, такую что для всех $x \in A$*

$$(\exists y \in B.R(x, y)) \Longrightarrow s(x) \downarrow \wedge R(x, s(x))$$

В классической теории вычислений есть хорошо известные «теоремы счетности», первая из них утверждает что «существует вычислимое перечисление вычислимо перечисляемых множеств», а вторая постулирует то же самое, но для частично вычислимых функций. В СТВ этим теоремам соответствуют следующая аксиома и утверждение 8. До сих пор, все что мы рассматривали вполне может существовать в классической логике (при соответствующем переходе к классической теории множеств), хотя и непредставляло бы тогда никакого интереса ($\mathbf{2} = \Sigma = \Omega$). Следующая аксиома в этом смысле является переломной – следствия из нее выходят за пределы классической логики.

Аксиома 5 (Счетности). *Множество счетных подмножеств* $\mathbb{N}$ *счетно.*

Следующее утверждение, является аналогом важной классической теоремы о вычислимости

Предложение 8. $\mathbb{N} \to \mathbb{N}$ *счетно*

Следующая теорема, которая является следствием аксиомы счетности, раскрывает суть возникающего в классической логике противоречия.

Теорема 7. *Ни одно из следующих включений не является равенством:*

$$\mathit{2} \subseteq \Sigma \subseteq \Omega_{\neg\neg} \subseteq \Omega$$

Таким образом, мы получаем «невозможное» неравенство $\mathbf{2} \neq \Omega$. И наконец $\mathbf{2} \neq \Sigma$ означает что семейства разрешимых и полуразрешимых подмножеств $\mathbb{N}$ не совпадают.

Предложение 9 (Принцип Фоа (Phoa)). *Для каждой* $f : \Sigma \to \Sigma$ *выполняются следующие два равенства:*

$$f(x) = (f(\bot) \vee x) \wedge f(\top) f(x) = f(\bot) \vee (x \wedge f(\top))$$

12. Фокальные множества

После подъема множество приобретает «неопределенный» элемент $\bot_A$ Но иногда, множество уже содержит такой «неопределенный» элемент. Например, он может быть получен путем присоединения $\top$ к множеству и затем отображением получившегося множества на исходное, не трогая элементы последнего. Таким образом, мы приходим к следующему определению.

Определение 22 (Фокальное множество). *Фокальным множеством* называется множество A взятое вместе с отображением (*фокусныи*) $\epsilon : A_\bot \to A$ при котором $\epsilon(\{x\}) = x$ для всех $x \in A$. Элемент $\epsilon(\bot_A)$ называется *точкой фокуса.*

Говорят, что множество A *связно* если оно не может быть разложено на несвязное объединение $A_1 + A_2$ нетривиальным способом. Другими словами в связном множестве любое отображение $A \to \mathbf{2}$ является константой.
Основываясь на этом определении в качестве заключения приведем две теоремы СТВ, являющиеся **более общими** аналогами теорем классической теории вычислений.

Теорема 8 (Райса). *Фокальное множество связно.*

Классическая теорема Райса утверждает, что не существует нетривиальных разрешимых подмножеств $\mathcal{E}$. Это сразу же следует из приведенной выше теоремы, т.к. $\mathcal{E}$ - фокальное множество.

Определение 23 (многозначная функция). *Многозначной функцией* $f : A \rightrightarrows B$ называется функция $f : A \to \mathcal{P}(B)$, такая что $f(x)$ невырождено для каждого $x \in A$.

Стационарной точкой многозначной функции $f : A \rightrightarrows A$ называется такая точка $x \in A$, что $x \in f(x)$.

В завершении, приведем обобщенную формулировку теоремы о рекурсии:

Теорема 9 (Теорема о рекурсии). *Каждая многозначная функция на счетном фокальном множестве имеет стационарную точку.*

Из теоремы 9 вытекает следующая классическая теорема.

Следствие 9.1 (Классическая теорема о рекурсии). *Для каждого* $f : \mathbb{N} \to \mathbb{N}$ *существует* $n \in \mathbb{N}$ *такое что* $\varphi_{f(n)} = \varphi_n$.

13. Заключение

На данном этапе знакомство с синтетической теории вычислений заканчивается. За рамками данной статьи остались описания пересечения СТВ с топологией, в частности, теоремы Райса-Шапиро и Шефердсона, описывающие топологии на $\Sigma^{\mathbb{N}}$ и $\mathbb{N} \to \mathbb{N}_{\perp}$ соответственно. Фактически, исследование рекурсивных топологий и рекурсивного анализа в рамках СТВ – это одно из направлений, напрашивающихся на разработку. Второе направление связано с изменением самого понятия вычисления и построение топосов, структура которых соответствовала бы данному понятию.

Небольшие достижения все же есть – формулировки и некоторые классические теоремы теории вычислимости выглядят в СТВ более элегантно. Доказаны аналоги основных теорем теории вычислимости (в некоторых случаях более общие), такие как теорема Райса, теорема о рекурсии,теорема Поста, теорема об однозначности, теорема Райса-Шапиро и другие. По мнению автора, СТВ еще не успела вырасти до актуальных проблем теории вычислимости, что, несомненно, не способствует привлечению к ней внимания со стороны «классических» специалистов. Тем не менее автор уверен, что применение синтетического подхода поспособствует получению новых результатов.

Литература

1. Цаленко, М.С. Основы теории категорий / М.С. Цаленко. – М.: Наука, 1974.
2. Голдблатт, Р. Топосы. Категорный анализ логики / Р. Голдблатт. – М.: Мир, 1983.

3. Маклейн, С. Категории для работающего математика / С. Маклейн. – М.: ФИЗМАТЛИТ, 2004.
4. Bauer, A. First steps in synthetic computability theory / A. Bauer // Electr. Notes Theor. Comput. Sci. – 2006. – V.155. – P.5-31.
5. Johnstone, P. Sketches of an Elefant. A Topos Theory Compendum / P. Johnstone. – Oxford: Clarendon Press, 2002.
6. Richman, F. Church's thesis without tears / F. Richman // The Journal of Symbolic Logic. – 1983. – V.48. – P.797-803.
7. Hyland, M. The effective topos / M. Hyland. – The L.E.J. Brouwer centerary symposium, 1982. – P.165-216.
8. Kock, A. Synthetic Differential Geometry / A. Kock. – Cambridge University Press, 2006.
9. Lavendhomme, R. Basic concepts of Synthetic Differential Geometry / R. Lavendhomme. – Kluwer Academic Publishers, 1996.
10. Phoa, W. An introduction to fibrations, topos theory, the effective topos and modest sets: Tech. rep. / W. Phoa. – The University of Edinburgh, 1992.
11. Rosolini, G. Continuity and effectiveness in topoi: Ph.D. thesis. – University of Oxford, 1986.
12. Rosolini, G. Domains in H. / G. Rosolini // Theoretical Computer Science, 2003. – V.2, N.264. – P.171-193.

Математические
структуры и моделирование
2009, вып. 19, с. 45–67

УДК 519:577.4+502.7

КАТАСТРОФЫ ТИПА «БАБОЧКА» В ЭВОЛЮЦИИ ЛЕСНЫХ ЭКОСИСТЕМ

Л.А. Володченкова, А.К. Гуц

Предлагается описание катастроф в экологии лесных биоценозов с помощью математической теории катастроф. Модель позволяет проводить качественное прогнозирование экологических катастроф и показывает к каким равновесным состоянием устремляется лесная экосистема после катастрофы.

Лес – это элемент географического ландшафта, состоящий из совокупности деревьев, занимающих доминирующее положение, кустарников, напочвенного покрова, животных и микроорганизмов, в своем развитии биологически взаимосвязанных, влияющих друг на друга и на внешнюю среду.

В этой статье ставится задача построения математической динамической модели смены равновесных состояний лесных экосистем, с чьей помощью можно было бы поводить прогнозирование экологических катастроф, при которых лесная экосистема теряет равновесие и устремляется к новому. Предложенная модель способна давать ответы на вопросы: каковы эти равновесия, когда они нарушаются, к каким равновесиям система перейдет и сколько новых равновесий может быть?

Модель основывается на четырех управляющих внешних факторах, задающих среду экосистемы. Это влажность почвы w; оконнная динамика u, определяющая мозаичность фитоценза; наличие конкуренции l и антропогенное вмешательства v в лесную экосистему (вырубка леса, пожары и т.д.). Модель может быть усложнена за счет введения дополнительных управляющих внешних факторов, но при этом она становится менее наглядной и требует при ее использовании уже гораздо более серьезных математических знаний.

1. Характеристики лесных экосистем

1.1. Факторы, влияющие на видовое разнообразие и богатство лесной экосистемы

1. Видовое богатство и видовая насыщенность зависят от многих факторов, в первую очередь от условий существования фитоценоза Главными лимитирую-

---

Омский государственный университет.
E-mail: `volodchenkova@cmm.univer.omsk.su`, `guts@omsu.ru`

щими факторами видового разнообразия являются температура, *влажность* и наличие пищевых ресурсов [1].

2. Условия произрастания могут сильно различаться по теплообеспеченности, *влажности* и богатству почв в пределах одного региона [1].

3. Видовое разнообразие взаимосвязано и с разнообразием условий конкретной среды обитания [1].

3. Сказывается и воздействие человека, и степень изменения среды самим сообществом (максимальная в лесах и минимальная в пустынях) [1].

1.2. Ярусность

Распределение растений по надземным *ярусам* определяет неодинаковая освещенность, которая приводит к различиям в температурном режиме и режиме влажности.

Каждый ярус, входящий в состав фитоценоза, оказывает влияние на другие ярусы и в свою очередь подвергается их влиянию. Поэтому фитоценоз необходимо рассматривать как нечто целое, а ярусы фитоценоза — как его структурные части, которые в некоторых случаях могут быть относительно самостоятельными.

Дюрье выделяет крупные ярусы: деревьев, кустарников, трав, напочвенного покрова, а в пределах этих ярусов — подъярусы:

– верхний, *первый ярус A*, образуют деревья. Его *подъярус A.1* состоит высоких деревьев; *подъярус A.2* – деревья второй величины – рябина обыкновенная, черёмуха обыкновенная, ива козья, дикая яблоня.

– *второй ярус леса B* состоит из кустарников, образующих подлесок – лещина обыкновенная, жимолость лесная, крушина ломкая, бересклет европейский.

– *третий ярус C* леса состоит из трав. Подъярус высоких трав – чистец лесной, бор развесистый, борцы; подъярус низких травам – сныть обыкновенная, осока волосистая, пролестник многолетний и др.

– *четвёртый ярус D* – мхи, грибы, лишайники.

Ярусное расположение растений связано с неодинаковой освещённостью. Количество света уменьшается от яруса к ярусу. Много света получают деревья первого яруса и очень мало – мхи и лишайники. В еловом лесу кустарники не растут – ветви елей задерживают очень много света, в таком лесу всегда сумрачно.

Ярусами располагаются и корни от лесных растений. Это позволяет поглощать воду, минеральные вещества из разных по глубине слоев земли.

«Каждый ярус, входящий в состав фитоценоза, оказывает влияние на другие ярусы и в свою очередь подвергается их влиянию. Поэтому фитоценоз необходимо рассматривать как нечто целое, а ярусы фитоценоза — как его структурные части, которые в некоторых случаях могут быть относительно самостоятельными» [1].

«Многоярусные биоценозы, представленные большим количеством видов растений, животных и микроорганизмов, связанных между собой разнообразными пищевыми и пространственными отношениями, называются сложными. Они наиболее устойчивы к неблагоприятным воздействиям. Исчезновение какого-либо вида существенно не отражается на судьбе таких биоценозов. В них происходит лишь незначительная перестройка организации, при которой популяции одного и даже нескольких видов могут заменяться экологически близкими видами, а стабильность сообщества определяется количественной регуляцией численности одних видов другими» [1].

1.3. Мозаичность. Оконная динамика

Обычно фитоценоз не бывает совершенно одинаковым на всем своем протяжении.

Горизонтальная неоднородность фитоценоза характеризуется *мозаичностью.*

«*Мозаичность* – горизонтальное расчленение внутри фитоценоза, обусловленное ценотическими факторами. Причины ее разные:

– неравномерность условий существования, вызванная жизнедеятельностью тех или иных растений: различия в затенении, химизме и физических особенностях опада, в нанорельефе;

– результат деятельности животных, выбрасывающих на поверхность почву из глубоких горизонтов (землерои) или нарушающих в некоторых участках обычный строй растительного сообщества (грызуны);

– способом роста тех или иных растений, образующих кочки или куртины. Так, высокие кочки, образуемые осокой дернистой на болотах, способствуют возникновению весьма разнообразных микроценозов на вершинках, на откосах и в понижениях между кочками» [1].

Основным механизмом смены поколений деревьев, поддержания специфической мозаичности лесов и высокого биологического разнообразия во многих типах леса является динамика, связанная с гибелью отдельных старых деревьев или их групп, образования за счет этого «окон» в пологе древостоя и формирования в этих окнах групп молодых деревьев. Этот механизм постепенной смены поколений деревьев получил название *оконной динамики* (а соответствующая ей мозаичность лесных экосистем – *оконной мозаики*).

Для лесов, где существует оконная динамика характерны следующие особенности:

– высокая мозаичность древостоя и всех остальных ярусов леса, связанная с постоянным вываливанием или (реже) усыханием «на корню» отдельных старых деревьев, их групп или целых участков древостоя и формированием на их месте новых, более молодых групп деревьев, а также преобладание или значительная доля участков с низкой сомкнутостью крон деревьев и высокой освещенностью нижних ярусов леса. Непрерывность процесса гибели крупных деревьев и образования окон обеспечивает поддержание разновозрастности древостоев, их стабильности во времени (при рассмотрении больших площадей леса) и сложной световой мозаики под пологом леса. Благодаря этой световой мозаике и наличию под пологом леса сильно освещенных участков, в таких лесах можно встретить как исключительно теневыносливые виды растений, например кислицу или седмичник, так и светолюбивые виды – малину, иван-чай и другие;

– непрерывный популяционный возрастной спектр большинства древесных пород, образующих древостой. В норме подобные леса являются полидоминантными с господством темнохвойных пород и примесью лиственных – в южной тайге ильма, липы, клена, дуба, ясеня, березы, в средней и северной тайге – березы. Как правило, темнохвойные породы деревьев (ель, пихта, сосна сибирская) характеризуется непрерывным возрастным спектром и наличием всех возрастных стадий и возрастов деревьев (всходы могут отсутствовать, что связано с неежегодным обильным образованием еловых семян). Лиственные породы деревьев могут не иметь непрерывного возрастного спектра, что связано как с меньшей их долей в составе древостоя по количеству особей (и, соответственно, большими флуктуациями численности), так и с большей их зависимостью от размера образующихся окон (и, соответственно, большей зависимостью от действия ураганов, экстремальных погодных условий, вспышек численности вредителей и болезней);

– наличие большого количества валежа разных размерных классов и разных степеней разложения, а также ветровального почвенного микрорельефа (включающего западины и бугры ветровально-почвенных комплексов разного размера и возраста). Валеж и ветровальный почвенный микрорельеф создают высокую степень неоднородности субстратных условий под пологом леса, обеспечивая, совместно со световой мозаикой, высочайшее разнообразие экологических ниш и возможность сосуществования в пределах одного участка леса огромного количества видов растений и животных;

– наличие мощных, насыщенных органическим веществом верхних почвенных горизонтов, обеспечивающее, вместе с валежом и ветровальным микрорельефом, высокую влагоудерживающую способность этих лесных экосистем, преобладание внутрипочвенного стока над поверхностным и поддержание особого микроклимата под пологом леса, характеризующегося повышенной влажностью воздуха и верхних почвенных горизонтов и, как следствием, сглаженными суточными колебаниями температуры;

– как правило, значительное участие или преобладание в составе травяно-кустарничкового яруса видов из группы таежного крупнотравья – щитовника австрийского, кочедыжника женского, диплазиума сибирского и других крупных папоротников, бодяков огородного и разнолистного, цицербиты альпийской или уральской и многих других сходных с ними по экологии видов растений.

Леса с оконной динамикой древесного яруса имеют ключевое значение для сохранения в естественных условиях большого количества видов таежной флоры и фауны. Кроме того, благодаря огромному количеству запасенного в почве этих лесов и на ее поверхности мертвого органического вещества, эти леса имеют очень большое водорегулирующее (как эффективная природная губка) и климаторегулирующее (как резервуары исключенного из атмосферного круговорота углерода) значение. К сожалению, таких лесов сохранилось исключительно мало, и с каждым годом их площадь сокращается все сильнее и сильнее – прежде всего, под воздействием вырубок и возникающих вокруг них и лесовозных дорог пожаров.

1.4. Атмосфера

«Рост и развитие растительности зависят от температуры, влажности воздуха, его движения и состава, но и наоборот – состав, высота, ярусность и густота растительности влияют на эти свойства атмосферы» [1].

1.5. Влажность почвы. Засуха. Избыток влаги

Для лесов опасность представлет не только засуха, нехватка влаги в почве, но и вымокание.

«Под вымоканием лесов понимается угнетение и гибель лесных массивов вследствие длительного застоя воды, заболачивания территории, подъема уровня грунтовых вод и засоления почв. Проблема вымокания лесов характерна для равнинных, слабо дренированных районов, с неглубоким залеганием грунтовых вод и повышенным увлажнением. В частности, она актуальна для обширных районов Западно-Сибирской равнины. В условиях повышенной обводненности грунтов и заболоченности, прокладка автомобильных дорог и магистральных нефте- и газопроводов требует возведения мощных насыпей, что приводит к нарушению гидрологического режима, и как результат – к деградации и вымоканию лесных массивов <..>

При вымокании происходят изменения в анатомии древесной растительности, в видовом составе травяного покрова, что в конечном итоге приводит к смене одной экосистемы на другую: на месте лесов возникают болота» [2].

1.6. Пожары

Пожары являются важным фактором, способным вывести лесную экосистемы из равновесия и способствовать развитию экологической катастрофы. При лесных пожарах повреждается или полностью уничтожается растущий лес вместе с подлеском, подростом и травяным покровом. В связи с этим утрачивается источник получения древесины и резко снижаются водоохранно-защитные и санитарно-гигиенические свойства леса. Пожары уничтожают гнезда птиц и местообитания зверей, способствуют размножению вредных насекомых.

«Лесные пожары очень существенно влияют на экологию лесов, включая формирование круговорота углерода. Пожары инициируют новую сукцессию полога леса и, таким образом, регулируют аккумуляцию углерода, определяемую первичной продукцией. Пожары воздействуют, помимо этого, на тепловой режим почвы, что, в свою очередь, оказывает влияние на процессы дыхания почвы. Наконец, следствием лесных пожаров оказывается их влияние на круговорот углерода в региональных и глобальных масштабах, обусловленное выбросами углерода в атмосферу и последующим дальним переносом <..>

С одной стороны, – лесные пожары – это стихия природного (а иногда и антропогенного) происхождения, причиняющая серьезный материальный ущерб. С другой стороны, – пожары – необходимый компонент эволюции лесов, обеспечивающий их обновление» [3].

1.7. Вырубка лесов

Вырубка леса без мер по его восстановлению означает обеслесивание и, следовательно, фактически наступающую экологическую катастрофу. К примеру, вырубка лесов резко изменяет характер стока и гидрологический режим рек, вызывает бурные весенние паводки и резкое обмеление рек в летнее время, водоохранные свойства леса снижаются или пропадают полностью.

Рубки в лесных экосистемах необходимо рассматривать как антропогенный фактор возножной потери равновесия экосистемой.

1.8. Принцип конкурентного исключения

Принцип конкурентного исключения гласит: если два вида конкурируют за одну экологичекую нишу[1], есть только два возможных исхода. Либо эти два вида немного изменятся и каждый займет немного другую нишу (дифференциация ниш), либо один из видов обречен на вымирание [2].

«Конкуренция отмечается между особями одного вида (внутривидовая борьба) и между особями разных видов (межвидовая борьба) в неблагоприятных условиях среды.

Оба типа борьбы обычно тесно связаны друг с другом. В борьбе за пищу, за влагу, с вредителями и паразитами растения конкурируют и с особями своего вида, и с особями других видов» [1].

1.9. Доброкачественность

Моделирование экологии лесных биоценозов, если мы желаем иметь численный показатель *качества лесной экосистемы* x, требует начать с определения этого показателя. Мы прежде всего предлагаем название для требуемого показателя, вводя термин *доброкачественность лесной экосистемы.*

Конкретное значение доброкачественности лесной экосистемы, находящейся в состоянии *равновесия*, будем интерпретировать как конретное растительное сообщество. Смена равновесия, сопровождающееся изменением значения доброкачественности лесной экосистемы, – это смена одних сообществ другими (динамика растительного покрова по Сукачёву).

Доброкачественность лесной экосистемы может быть охарактеризована достаточно большим числом количественных показателей, которые получаются при проведении различных замеров, производимых в лесной экосистеме.

[1] Под экологической нишей подразумевают совокупность факторов окружающей среды, в пределах которых данный вид может развиваться и воспроизводиться.

[2] См. сайт http://elementy.ru

Перечислим некоторые из них [1]:

1. *Бонитет*[3] леса – показатель продуктивности леса, зависящий от почвенно-грунтовых и климатических условий (местообитания). Определяется средней высотой деревьев господствующей породы насаждения с учётом его возраста. В советской таксационной практике пользуются шкалой классов бонитета, составленной в 1911 проф. М.М.Орловым. По бонитировочной шкале насаждения делятся на 5 классов бонитета, обозначаемых римскими цифрами. К I классу относят насаждения наиболее продуктивные, к V классу – наименее продуктивные. Нередко число классов бонитет леса увеличивают, например знаком Ia обозначают насаждения с продуктивностью выше I класса и знаком Va – ниже V класса. Для всех древесных пород принята общая бонитировочная шкала. Для семенных и порослевых насаждений установлены особые шкалы;

2. Показатели *численности видов* и их динамика являются одними из основных в изучении растительности. Численность определяется визуально и инструментально, но чаще визуально. Всегда на учетной единице: площади (дм, м2, км2, га), длины (м, км), объема (м3, 10 дм3), времени (час, сутки) и т.д. [1]

3. Количественное соотношение видов в биоценозе, или *индекс разнообразия* H чаще всего определяется по формуле Шеннона:

$$H = -\sum_i p_i \log p_i,$$

где p_i – доля вида в сообществе [1].

4. *Встречаемость* (частота встречаемости, коэффициент встречаемости) – это относительное число выборок, в которых встречается вид. Если выборка состоит из 100 учетных площадок, а вид отмечен на 43, то и встречаемость равна 43%. При встречаемости 25%, вид встречается в каждой четвертой площадке учета и он случайный. Высокая встречаемость, если вид отмечен более, чем на 50% уч. пл. Обычно закладывается 50 уч. пл., но не менее 25. [1]

5. *Обилие* – это количество особей вида на единице площади или объема. Наиболее часто используются шкалы обилия Друде и Хульта.

6. *Покрытие* – процент площади, покрываемой надземными частями растений. Процент площади, занятой основаниями растений -– истинное покрытие, верхними частями — проективное. Проективное покрытие -– обязательный показатель при изучении напочвенного покрова.

7. *Биомасса* – общие запасы органического вещества, накопленные к моменту учета. Выражаются в массе абсолютно-сухого, воздушно-сухого или сырого вещества. Биомасса растений — растительная масса, фитомасса; биомасса животных – зоомасса. Биомасса, ее фракционная структура, скорость накопления (продукция – прирост биомассы за определенный промежуток времени) являются важнейшими – интегральными, показателями жизнедеятельности организмов. Они дают возможность оценить роль каждого фактора и популяции в формировании биогеоценоза, оценить запасы биологических и пищевых ресурсов, сделать кратко- и долгосрочные прогнозы развития сообществ, предсказать

[3] Bonitas (лат.) – доброкачественность.

пути их трансформации и разработать мероприятия по охране и рациональному использованию любого из ресурсов. Именно поэтому изучение биологической продуктивности и было положено в основу упомянутой в одной из предыдущих лекций Международной биологической программы (МБП).

8. *Обилие вида* – число особей данного вида на единицу площади или объема занимаемого ими пространства.

9. *Степень доминирования* – отношение (обычно в процентах) числа особей данного вида к общему числу всех особей рассматриваемой группировки. Степень доминирования – отношение (обычно в процентах) числа особей данного вида к общему числу всех особей рассматриваемой группировки.

К этим характеристикам можно добавить и другие:

10. *Продуктивность древостоя* – количество стволовой древисины, коры, сучьев, ветвей, листьев, хвой и корней древостоя на единицу площади.

11. *Санитарное состояние леса* – характеристика леса, содержащая сведения о его захламленности, наличии усыхающих и сухостойных деревьев.

12. *Суховершинность* – наличие сухой вершины у растущего дерева.

13. *Влажность древесины* – отношение массы воды, содержащейся в древесине, к массе древесины, в процентах.

15. Обстоятельство, определяющее видовую насыщенность и видовое богатство: *возраст данного фитоценоза*, т. е. продолжительность времени, в течение которого растения, входящие в состав фитоценоза, сживались друг с другом

2. Модель доброкачественностного леса

В § 1.9 перечислены всевозможные показатели x_j, $j = 1, 2, ..., 15$, которые могут быть взяты как показатель доброкачественности лесной экосистемы.

Однако лучше всего ввести интегральный показатель доброкачественности, имеющий вид

$$x = \sum_{j=1}^{15} w_j x_j,$$

где w_j – вес показателя x_j, т.е. его вклад (доля) в интегральный показатель. Значения показателя x в момент времени t обозначаем как $x(t)$. Это число принимается нами как степень доброкачественности лесного биоценоза.

Очевидно, что такой подход является крайне упрощенным, но любая модель есть определенное упрощение и со временем, естественно, может быть усложнена.

Степень доброкачественности лесного биоценоза в момент времени t будет определять доброкачественность леса в следующий момент времени $t + dt$.

Иначе говоря,

$$x(t + dt) = x(t) + A(t, x)dt, \tag{1}$$

где $A(t, x)dt$ – величина, описываюшая отклонения в степени доброкачественности леса, произошедшие на отрезке времени dt.

Из (1) имеем то, что называется дифференциальным уравнением

$$\frac{dx}{dt} = A(t, x). \tag{2}$$

В основе продуктивного процесса растений лежит фотосинтез. Растения под воздействием солнечной энергии, поглощая листьями из атмосферы углекислый газ и корневой системой из почвы *воду*, создают органическое вещество. Недостаток влаги $(-w) > 0$ является фактором не способствуюшим благополучию леса.

В таком случае следует в правую часть дифференциального уравнения (2) добавить член $-(-w)$:

$$\frac{dx}{dt} = A(t, x) - (-w). \tag{3}$$

В самом простом случае можно принять, что $A(t, x) = k_1 x$, т.е.

$$\frac{dx}{dt} = k_1 x + w. \tag{4}$$

Коэффициент k_1 можно посчитать постоянным. Но тогда доброкачественность леса будет нарастать как геометрическая прогрессия и это делает бессмысленной нашу модель. Поэтому начнем ее усложнять.

Коэффициент k_1 отвечает за «прирост доброкачественности леса». Учтём, что для здорового леса обязательной чертой является наличие ярусности.

Следуя Дюрье примем в расчет только четыре яруса. Взаимодействие ярусов должно дать вклад в доброкачественность леса вида

$$\alpha_1 x \cdot \alpha_2 x \cdot \alpha_3 x \cdot \alpha_4 x = \alpha_1 \alpha_2 \alpha_3 \alpha_4 x^4.$$

Каждый коэффициент $\alpha_i > 0$ характеризует степень участия i-го яруса во взаимодействии ярусов. Если $\alpha_i \to 0$, то мы констатируем отсутствие i-го яруса; означающее меньшую ярусность биоценоза и, следовательно, его меньшую степенть биоразнообразия, меньшую устойчивость к неблагоприятным воздействиям.

Примем. что

$$k_1 = k_0 x^4 - p(x), \tag{5}$$

где $k_0 = \alpha_1 \alpha_2 \alpha_3 \alpha_4 = const$ – «сила» взаимодействия четырех ярусов леса, а величина $p(x)$ – это то, что мешает доброкачественности леса.

Тогда следует принять, что

$$p(x) = \underbrace{k_2}_{\text{вырубка леса, пожары}} + \underbrace{k_3 x^2}_{\text{конкуренция}} - \underbrace{k_4 x}_{\substack{\text{оконная динамика} \\ \text{(хорошая атмосфера)}}} \tag{6}$$

где член $k_3 x^2$ храктеризует конкуренцию растений в лесу, а член $k_4 x$ допускает две интерпретации. При $k_4 > 0$, при первой интерпретации, имеем хороший состав атмосферы, или, при второй – действенен основной механизм смены поколений деревьев, поддерживающий специфическую мозаичность лесов и высокое биологическое разнообразие во многих типах леса, называемый «оконной динамикой». Он связан с гибелью отдельных старых деревьев или их групп, образования за счет этого «окон» в пологе древостоя и формирования в этих окнах групп молодых деревьев.

Объединяя уравнения (4)-(6), получаем

$$\frac{dx}{dt} = k_0x^5 - k_2x - k_3x^3 + k_4x^2 + w =$$

$$= \frac{\partial}{\partial x}\left\{\frac{k_0}{6}x^6 + \frac{-k_3}{4}x^4 + \frac{k_4}{3}x^3 + \frac{-k_2}{2}x^2 + wx\right\} \qquad (7)$$

или

$$\frac{dx}{dt} = \frac{\partial}{\partial x}V(x,l,u,v,w), \qquad (8)$$

где

$$V(x,l,u,v,w) = \frac{k_0}{6}x^6 + lx^4 + ux^3 + vx^2 + wx, \qquad (9)$$

$$l = -\frac{k_3}{4}, \quad u = \frac{k_4}{3}, \quad v = -\frac{k_2}{2}.$$

Отметим, что доброкачественность характеризуется неравенством $x > 0$; наличие конкуренции неравенством $l < 0$, действенность оконной динамики неравенством $u > 0$; вырубка лесов, пожары неравенством $v < 0$, недостаток влаги неравенством $w < 0$.

Функция $V(x,l,u,v,w)$, заданная выражением (9), описывает при изменении параметров l, u, v, w самые различные бифуркации, называемые в математической теории катастроф катастрофами типа «бабочка» [5, 6, 6, 7].

3. Равновесия системы, их смена и математическая теория катастроф

Равновесные системы, и это относится и к лесным экосистемам, характеризуются тем, что их строение и состав колеблются около какой-то средней точки, представляющей как бы типичное состояние растительного покрова.

Равновесие в природе на самом деле зависит от окружающей среды, а среда эта постоянно подвержена изменениям. Пожары, наводнения, колебания количества атмосферных осадков оказывают влияние на среду, в которой произрастает лес. И растения, конечно же, не могут не реагировать на эти изменения. Получается, что экосистема все время пытается либо сохранить равновесие, либо попасть в новое равновесие. Вмешательство человека – всего лишь еще один способ изменить окружающую среду и, таким образом, повлиять на направление развития экосистемы.

Экосистему можно вывести из состояния равновесия многими способами. Обычно это бывает пожар, наводнение или засуха. После такого нарушения равновесия новая экосистема сама себя восстанавливает, и этот процесс носит регулярный характер, называется *сукцессией*, и повторяется в самых разных ситуациях.

Сукцессия – это последовательный ряд смены серийных (временно существующих) растительных сообществ на конкретном местообитании после выведения конкретной экосистемы из состояния равновесия [1]. «При этом конкретная экосистема возвращается в свое исходное состояние и пребывает в нем до тех пор, пока не изменятся климат, рельеф, гидрологический режим, пока вновь не пройдет пожар, или не случится какая-то другая катастрофа. И вновь начнется новая сукцессия, которая либо приведет к восстановлению исходного сообщества, либо нет. В результате сукцессии на конкретном местообитании восстанавливается исходное растительное сообщество, называемое геоботаниками климаксовым, или коренным. Коренное сообщество растений устойчиво и в данных климатических условиях не изменяется» [1].

В данной работе не ставится задача описания сукцессии, т.е. задача возврата к исходному равновесию, а описывается выведение экосистемы из равновесия и прогнозируются возможные *новые* равновесия, к которым преходит экосистема.

Для математического описания смены равновесий системы при изменении внешних (управляющих) факторов математиками создана теория элементарных катастроф. Она исходит из предположения, что поведение системы определяется некоторой потенциальной функцией $V = V(x, u_1, ..., u_n)$, где x – переменная, характеризующая изучаемую систему, а $u_1, ..., u_n$ – внешние управляющие факторы.

Динамика изменения во времени переменной $x(t)$ задается дифференциальным уравнением

$$\frac{dx}{dt} = \frac{\partial V(x, u_1, ..., u_n)}{\partial x} \qquad (10)$$

Равновесия системы – это такие решения $x(t)$ уравнения (1), которые не меняются со сременем (какой-то период времени), т.е.

$$\frac{dx}{dt} = 0.$$

Но в таком случае, как видно из уравнения (1), имеем *уравнение всевозможных равновесий* системы

$$\frac{\partial V(x, u_1, ..., u_n)}{\partial x} = 0. \qquad (11)$$

Каждое равновесие есть решение x данного уравнения, зависящее от конкретного набора внешних управляющих факторов, т.е. $x = x(u_1, ..., u_n)$.

Если меняются внешние факторы, скажем вместо набора $(u_1^0, ..., u_n^0)$ переходим к набору $(u_1^1, ..., u_n^1)$, то имеем смену равновесия системы – вместо равновесия $x_0 = x(u_1, ..., u_n)$ получаем равновесие $x_1 = x(u_1^1, ..., u_n^1)$.

Естественно считать, что если слабо изменятся внешние факторы, то система почти не изменит своего состояния, т.е. практически сохранит значение характеризующей ее пременной x_0, или займет равновесие близкое к тому, что было, конечно же не отличающееся от предыдущего качественно. Однако исследования показали, что возможны такие малые изменения внешних факторов, и они происходят тогда, когда они пересекают при своем изменении так называемые *бифуркационные множества*, при которых система переходит к новому равновесию, для которого ее характерстика x_1 существенно отлична от предыдущей и, следовательно, новое равновесие обладает новыми качествами.

Такие переходы были названы катастрофами, поскольку переходы к новому *устойчивому* равновесию предшествует потеря устойчивости раннего равновесия. Название было продиктовано примерами из механики, физики и кораблестроения, где такие процессы описывали реальные катастрофы (перелом балки, замерзание воды, перевертывание судна). Отсюда и название математической теории, данное ей французским математиком Рене Томом.

С точки зрения математики, как видно из уравнения (11), равновесие $x = x(u_1, ..., u_n)$ – это либо *точка минимума*, либо *точка максимума*, либо так называемая *точка перегиба* функции $V = V_{(u_1,...,u_n)}(x) = V(x, u_1, ..., u_n)$. Обозначение $V_{(u_1,...,u_n)}(x)$ говорит о том, что мы функцию $V = V(x, u_1, ..., u_n)$ рассматриваем как функцию *одной* переменной x и используем соответствующий математический аппарат, известный школьникам.

Как правило, устойчивые равновесия системы – это минимумы функции $V = V_{(u_1,...,u_n)}(x)$. На рисунках графика функции $V = V_{(u_1,...,u_n)}(x)$ они изображаются ямками (соответственно максимумы – неустойчивые равновесия – изображаются вершинами горок).

Теория элементарных катастроф по самой своей природе локальна. Иначе говоря, функция $V_{(u,v,w)}(x) = V(x, u, v, w)$ как функция x определена лишь в окрестности $(-\varepsilon, +\varepsilon)$, и, следовательно, теория не описывает все возможные состояния равновесия. Уход системы в такие неописываемые теорией равновесия оговаривается как переход к равновесию $x = -\infty$ (минус бесконечность).

4. Определения экологической ситуации и экологических катастроф

Катастрофы, описываемые теорией катастроф вполне отвечают тому, что можно было бы назвать экологическими катастрофами[4]. Поэтому адаптируем терминологию теории катастроф к экологии.

Назовем *местной экологической ситуацией*, отвечающей набору внешних факторов $(u_1, ..., u_n)$ любую из точек локального минимума функции $V_{(u_1,...,u_n)}(x)$.

Экологическая система в случае местной экологической ситуации пребывает в *равновесии*, т.е. величина $x(t)$ не меняется со временем (управляющие факторы $(u_1, ..., u_n)$ зафиксированы).

Региональной экологической ситуацией, связанной с V, называется набор местных экологическая ситуация, предполагая и возможность равновесия $x = -\infty$ (минус бесконечность).

Принятие значение $x = -\infty$ означает, что наша математическая модель говорит о том, что экосистема переходит в равновесное состояние, которое не

[4]В общем смысле экологическая катастрофа означает достаточно быстрый переход системы к новому качественному равновесному состоянию.

описывается с помощью предложенного уравнения ее эволюции (1) (теория локальна!).

Фактически региональная экологическая ситуация – это набор местных экологических ситуаций.

Точкой экологической катастрофы называется любая точка перегиба функции $V = V_{(u_1,...,u_n)}(x)$.

Морфологией катастрофы или множеством катастроф называется множество всех точек катастрофы.

Точки катастроф, которые входят в бифуркационное множество, находят, исключая x, в процессе решения системы уравнений

$$\frac{\partial V(x,u_1,...,u_n)}{\partial x} = 0, \quad \frac{\partial^2 V(x,u_1,...,u_n)}{\partial x^2} = 0.$$

Для конкретного набора внешних факторов у функции $V = V_{(u_1,...,u_n)}(x)$ может быть несколько точек минимумов. Скажем это точки $x_1, x_2, ..., x_m$. В каком из этих равновесий находится система? Для этого придуманы различные *правила.*

Правилом называется способ, по которому мы выбираем равновесия, т.е. минимумы функции $V = V_{(u_1,...,u_n)}(x)$ для некоторой региональной экологической ситуации, связанной с V.

Рассмотрим два основных правила.

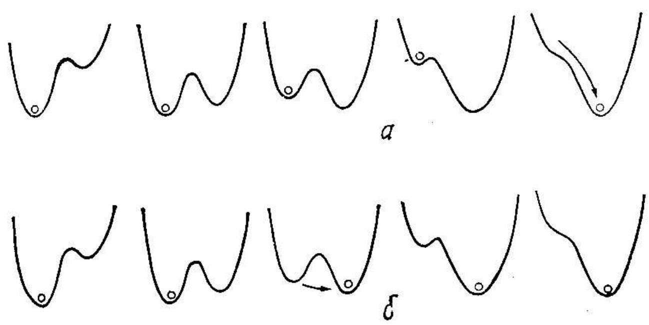


Рис. 1. Вверху показано действие правила максимального промедления, а на нижних рисунках – правило Максвелла.

Правило максимального промедления предписывает состоянию оставаться в минимуме при заменении факторов до тех пор, пока он существует (рис.1, вверху).

В момент, когда происходит исчезновение старого миниума и становится необходимым переход в новый, происходит экологическая катастрофа.

Правило Максвелла предписывает взять в качестве равновесия системы такую точку минимума, в которой функция $V = V_{(u_1,...,u_n)}(x)$ достигает наименьшего значения (рис.1, внизу).

Поскольку одним из значений этого минимума может оказаться $-\infty$, правилом Максвелла лучше пользоваться тогда, когда V имеет конечные минимумы.

Ясно, что катастрофы возникают тогда, когда $V = V_{(u_1,\ldots,u_n)}(x)$ достигает абсолютного минимума в двух различных местах.

В случае правила максимального промедления катастрофе, т.е. резкой смене равновесия отвечает качественное изменение формы потенциальной функции $V = V_{(u_1,\ldots,u_n)}(x)$. В случае правила Максвелла резкая смена равновесия происходит без качественного изменения формы функции $V = V_{(u_1,\ldots,u_n)}(x)$. С точки зрения теории катастроф в случае правила Максвелла можно говорить о некатастрофической смене равновесия. Хотя с точки зрения экологии это скорее всего самая настоящая катастрофа. Катастрофы происходят в любом смысле и вне зависимости от принятого правила, если, изменяясь управляющие внешние факторы пересекают бифуркационное множество.

Изучим все сказанное в нашем конкретном случае, когда функция V имеет вид (9).

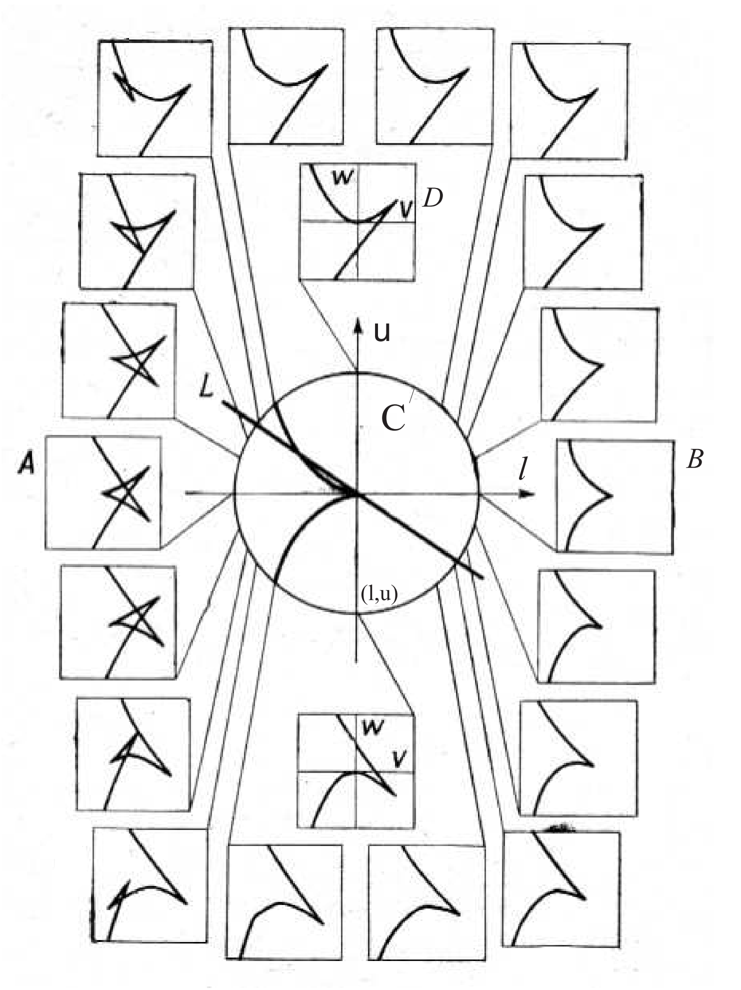


Рис. 2. Части бифуркационного множества B_V, лежащие в плоскостях vw, когда фиксируется точка $(l, u) \in C$ на окружности C. Кривая с точкой возврата в плоскости lu – это линия ласточкиных хвостов, т.е. геометрическое место точек, в которых происходит катастрофа ласточкин хвост [6].

5. Катастрофы в экологии леса типа «бабочка»

Для катастрофы «бабочка» потенциал

$$V(x,u,v) = x^6 + lx^4 + ux^3 + vx^2 + wx.$$

Рассмотрим множество

$$M_V = \{(x,l,u,v,w) : \frac{\partial}{\partial x}V = 6x^5 + 4lx^3 + 3ux^2 + 2vx + w = 0\}.$$

Оно состоит из максимумом, минимумов и точек перегиба функции $V_{(l,u,v,w)}(x)$. Все они отвечают равновесию в состоянии изучаемой лесной экосистемы.

Множество

$$S_V = \{(x,l,u,v,w) \in M_V : \frac{\partial^2}{\partial x^2}V = 30x^4 + 12lx^2 + 6ux + 2v = 0\}$$

состоит из точек перегиба, которые принимаются, когда точка (l,u,v,w) принадлежит так называемому *бифуркационному множеству*:

$$B_V = \{(l,u,v,w) \in prS_V : \exists x(6x^5 + 4lx^3 + 3ux^2 + 2vx + w = 0 \;\&$$

$$\& \; 30x^4 + 12lx^2 + 6ux + 2v = 0\}.$$

Это множество изображается в качестве поверхности в четырехмерном пространстве переменных l,u,v,w. Различные его сечения в плоскости vw при фиксированных l,u даны на рис.2.

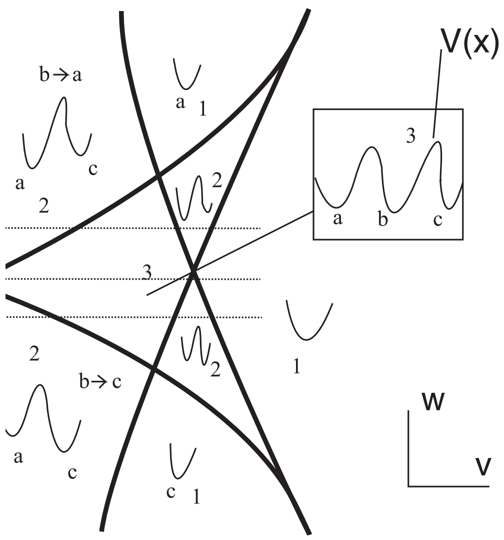


Рис. 3. Катастрофические смены экологических ситуаций при фиксированных $l < 0, u = 0$ в случае картинки A из рис.2. Дана динамика смены равновесий у функции $V_{(l,u,v,w)}(x)$ в сечении $l < 0, u = 0$. При переходе через линии B_V на плоскости vw рождается или умирает равновесие.

5.1. Катастрофы при вырубке лесов, пожарах. Нет мозаичности, есть сильная кокуренция. Влажность любая

Пусть дана ситуация, данная на A из рис.3. Тогда $u = 0$ (нет мозаичности леса), $l < 0$ (есть сильная конкуренции). Влажность почвы любая.

Рассмотрим случай, когда $w = const$, а внешний фактор v уменьшается от значения $v > 0$ (поддержание леса в порядке) через $v = 0$ до $v < 0$ (идет вырубка лесов, пожары) вдоль горизонтальных линий на рис.3, пересекающих область с цифрой 3. Тогда единственное устойчивое равновесие сменяется одним из трех возможных; в свою очередь, одно из них вскоре исчезает при дальнейшем уменьшении v, и лесная экосистема, наконец-то перейдет к новому равновесию, причем либо одному из двух альтернативных, либо к одному конкретному в зависимости от степени влажности почвы (рис.3).

5.2. Катастрофы при вырубке лесов, пожарах. Нет мозаичности, нет сильной конкуренции. Влажность любая

Пусть дана ситуация, данная на B из рис.4. Тогда $u = 0$ (нет мозаичности леса), $l > 0$ (отсутствует сильная конкуренции). Влажность почвы любая. Рассмотрим случай, когда $w = const$, а внешний фактор v уменьшается от значения $v > 0$ (поддержание леса в порядке) через $v = 0$ до $v < 0$ (идет вырубка лесов, пожары) вдоль горизонтальных линий 1,2,3 на рис.4, пересекающих линию B_V. Тогда единственное устойчивое равновесие сменяется одним из двух возможных и лесная экосистема, наконец-то перейдет к новому равновесию, причем либо одному из двух альтернативных, либо к одному конкретному в зависимости от степени влажности почвы (рис.4).

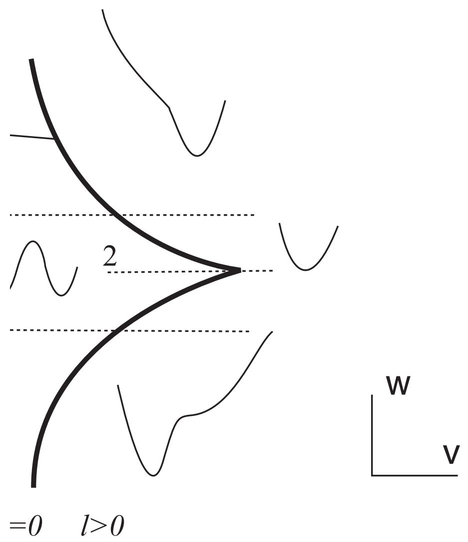


Рис. 4. Катастрофические смены экологических ситуаций при фиксированных $l > 0, u = 0$ в случае картинки B из рис.2. Дана динамика исчезновения минимумов у функции $V_{(l,u,v,w)}(x)$. Возможно не более двух равновесий.

5.3. Катастрофы при вырубке лесов, пожарах. Есть чистая хорошая атмосфера, нет сильной конкуренции. Влажность любая

Дана ситуация D из рис.5. Тогда $u > 0$ (чистая хорошая атмосфера), $l = 0$ (отсутствует сильна конкуренции). Влажность почвы любая. Рассмотрим случай, когда $w = const$, а внешний фактор v уменьшается от значения $v > 0$ (поддержание леса в порядке) через $v = 0$ до $v < 0$ (идет вырубка лесов, пожары) вдоль горизонтальных линий на рис.5, пересекающих линию B_V, например по оси v. Тогда единственное устойчивое равновесие сменяется одним из двух возможных и лесная экосистема, наконец-то перейдет к новому равновесию, причем либо одному из двух альтернативных, либо к одному конкретному в зависимости от степени влажности почвы (рис.5).

Если же $v = const < 0$ (пожары) и идет засуха (w уменьшается от 0 до очень маленького $w < 0$ вдоль линии 1 (рис.5), то два альтернативных равновесия сменятюся одним – лес после засухи и пожаров.

Если $v = const > 0$ (нет пожаров) и начинается и нарастает засуха (w уменьшается от $w > 0$ до очень маленького $w < 0$ вдоль линии 2 (рис.5), то одно равновесие заменяется на одно из двух альтернативных, а затем остается только одно – лес после засухи без пожаров.

Если $v = const > 0$ (нет пожаров) и начинается и нарастает засуха (w уменьшается от $w > 0$ до очень маленького $w < 0$ вдоль линии 3 (рис.5), то равновесие остается единственным, меняется только доброкачественность леса.

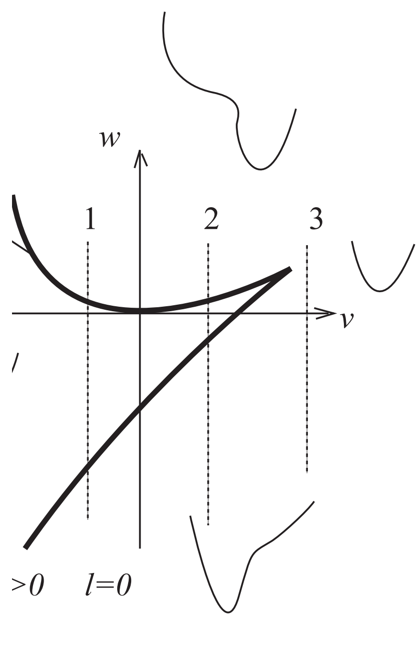


Рис. 5. Катастрофические смены экологических ситуаций при фиксированных $l = 0, u > 0$ в случае картинки D из рис.2. Дана динамика исчезновения минимумов у функции $V_{(l,u,v,w)}(x)$. Возможно не более двух равновесий.

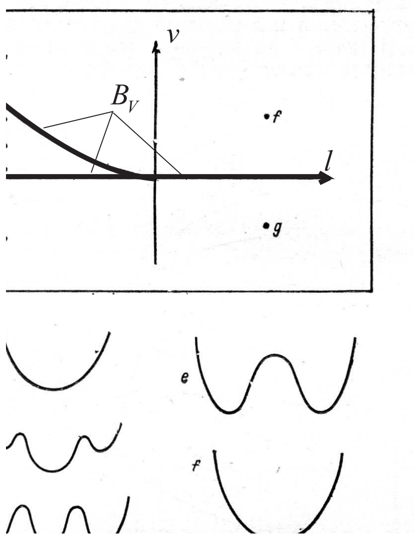


Рис. 6. Катастрофическая смена равновесий в случае потенциала $V_{(l,0,v,0)}(x)$ при изменении фактора вырубки лесов (пожаров). Если конкуренция действенна ($l < 0$), то вырубка леса (пожары), т.е. уменьшение v от $V > 0$ к $v < 0$ заставляет единственное устойчивое равновесие скачком поменять на одно из трех возможных в зависимости от принятого правила (a,b,c,d). При отсутствии кокуренции ситуация принципиально не поменяется (e,f,g) (рис. из [6, с.197]).

5.4. Катастрофы вырубки лесов и пожаров. Влажность $w = 0$

Пусть $u = w = 0$, т.е. не действенна оконная динамика (состояние атмосферы нейтральное) и нормальная влажность почвы.

Тогда мы будем наблюдать катастрофическую смену равновесий в случае потенциала $V_{(l,0,v,0)}(x)$ при изменении фактора вырубки лесов (пожаров).

Если конкуренция действенна ($l < 0$), то вырубка леса (пожары), т.е. уменьшение v от $v > 0$ к $v < 0$ заставляет единственное устойчивое равновесие скачком поменять на одно из трех возможных в зависимости от принятого правила (a,b,c,d) (см. рис.6).

При отсутствии кокуренции ситуация принципиально не поменяется (e,f,g) (см. рис.6).

5.5. Правило Максвелла: некатастрофическая смена равновесия при вымокании леса

Пусть имеет место случай $u = 0$, т.е. когда не наблюдается действие оконной динамики.

Рассмотрим четыре прямые 1,2,3, 4, заданные соответсвенно уравнениями:

$$1 = \{l > 0, u = 0, v = 0, w - \text{меняется}\},$$

$$i = \{l_i < 0, u = 0, v_i = (3/2)l_i^2, w - \text{меняется}\}, \quad (i = 2, 3, 4).$$

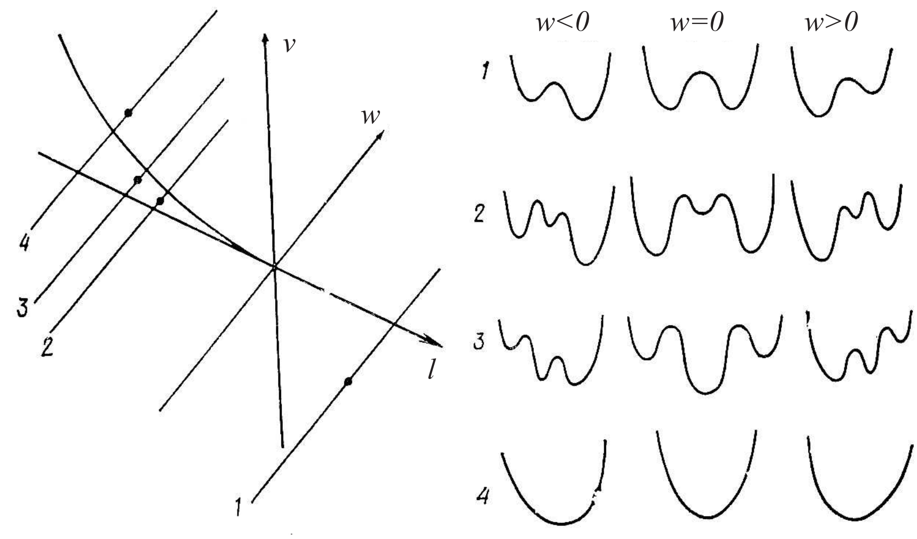


Рис. 7. Дана смены равновесий в случае потенциала $V_{(l,0,v,w)}(x)$ при возрастании влаги (избыток влаги, вымокание леса) в почве леса w. По правилу Макселла экосистема должна переходит к новому равновесию с меньшей доброкачественностью (рис. из [6, с.201]).

Вдоль прямых 1,2,3,4 лес поддержается в порядке (нет вырубки, пожаров), поскольку $v > 0$. См. рис.7 (слева).

Мы видим, что при возрастании избытка влаги в почве потенциальная функцция $V_{(l,0,0,w)}(x)$ экосистемы начинает меняться так, что при использовании нами правила Максвелла происходит смена равновесия и экосистема занимает состояние с меньшей доброкачественностью леса (рис.7, справа).

Если, однако, пользоваться правилом максимального промедления, то смены равновесия мы не наблюдаем. Такая неадекватность правил реалиям эволюции экологических систем демонстрирует скорее определенную незавершенность теории катастроф как совершенной математической теории, чем ее непригодность.

6. Катастрофы в плоскости «атмосфера (освещенность) u – антропогенный фактор) v»

На рис.2 приведена структура бифуркационного множества в плоскости (v, w) в зависимости от выбора конкретных значений конкуренции l и антропогенного фактора u.

На рис.8, 9 показана структура бифуркационого множества в плоскости «атмосфера (освещенность) u – пожары (вырубка) v» в зависимости от выбора значений конкуренции l и влажности w.

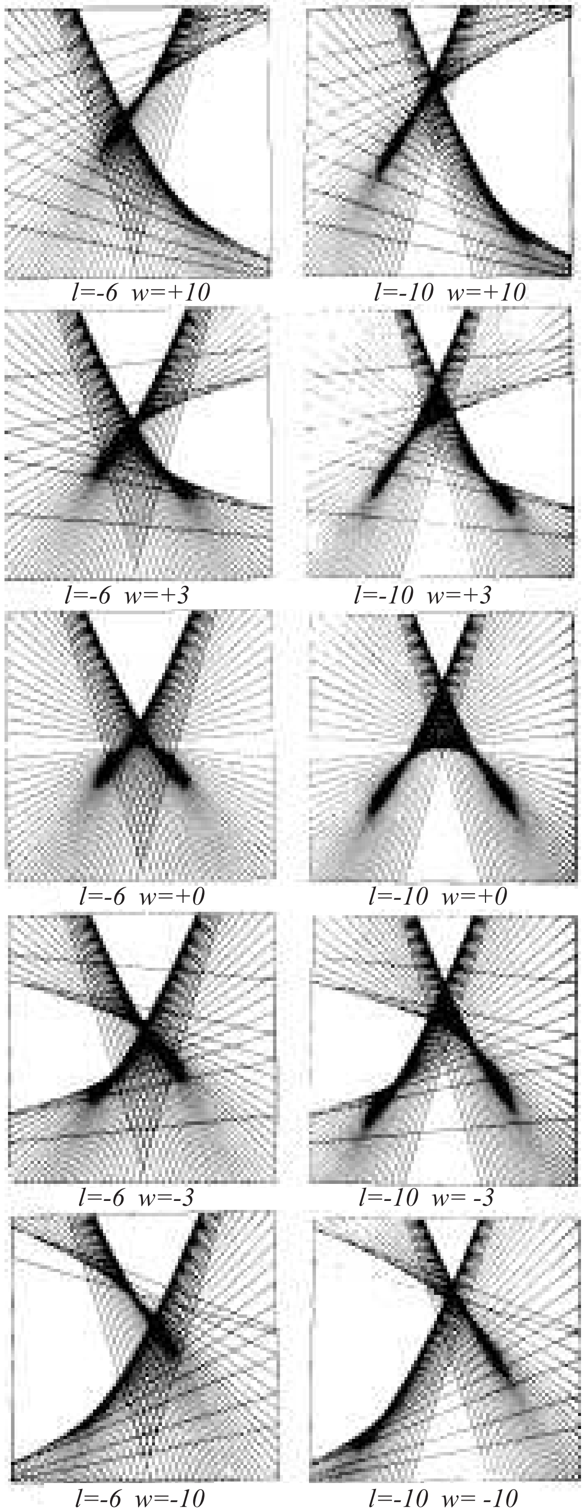


Рис. 8. Бифуркациооное множество в плоскости «атмосфера (освещенность) u – пожары (вырубка) v» в зависимости от значения конкуренции l (от сильной конкуренции $l = -6, -10$) и влажности w. Линии, принадлежащие бифуркационному множества в плоскости (u, v), – это огибающие семейств прямых линий, изображенных на рисунке [9, p.24].

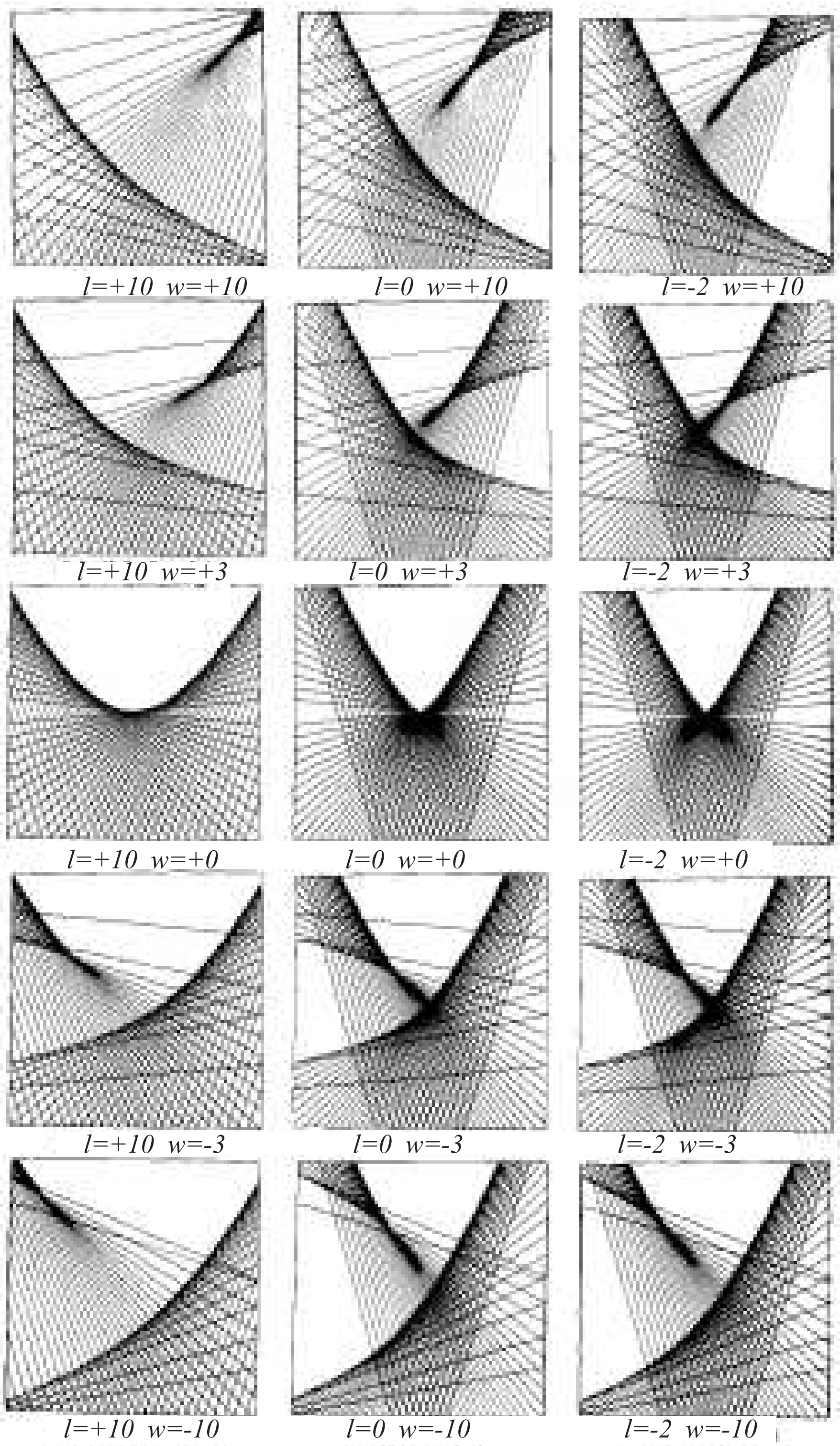


Рис. 9. Бифуркациооное множество в плоскости «атмосфера (освещенность) u – пожары (вырубка) v» в зависимости от значения конкуренции l (от отсутствия конкуренции $l = +10$ до ее наличия $l = -2$) и влажности w. Линии, принадлежащие бифуркационному множества в плоскости (u, v), – это огибающие семейств прямых линий, изображенных на рисунке [9, p.24].

7. Структурная устойчивость модели

Описание экологии леса, предложенное в данной статье, обладает важным достоинством – модель структурно устойчива. Это означает, что найденная нами потенциальная функция (9)

$$V(x,l,u,v,w) = \frac{k_0}{6}x^6 + lx^4 + ux^3 + vx^2 + wx, \tag{12}$$

в случае ее возмущения (малого шевеления) вида

$$V(x,l,u,v,w) \to V(x,l,u,v,w) + \delta V(x,l,u,v,w)$$

может быть возвращена к выражению внешне такому же[5], как (12), если совершить следующие преобразования координаты x и параметров l,u,v,w:

$$x \to \overline{x} = \overline{x}(x,l,u,v,w),$$

$$(l,u,v,w) \to (\bar{l},\overline{u},\overline{v},\overline{w}) = \phi(l,u,v,w)$$

и прибавить к возмущенному потенциалу некоторую функцию $\psi(l,u,v,w)$ [6, 6, 7].

Иначе говоря, наша модель устойчива по отношению к изменениям потенциальной функции, которые могут проявляться в форме пожеланий внести уточнения в вид потенциальной функции, или в форме высказываний о том, что в реальности лес не может быть описан некоторой выбранной раз и навсегда конкретной математической формулой.

Наличие структурной устойчивости у модели говорит о том, что ничего качественно нового в описании экологических ситуаций и экологических катастроф возмущенная модель не дает.

Наконец, если бы мы заявили о том, что нужно учесть в описании леса еще какой-либо один подъярус как ярус, то имели бы вместо (12) функцию

$$V(x,l,u,v,w) = \frac{k_0}{7}x^7 + lx^4 + ux^3 + vx^2 + wx, \tag{13}$$

которая не обладает свойством структурной устойчивости. Это, с одной стороны, говорит о том, что модель с выбором четырех ярусов не случайна, а с другой – что учёт подъярусов требует достаточно радикального изменения всей модели за счёт добавления новых управляющих параметров.

Тем не менее, забыть о структурной устойчивости заставляет то, что часто при выделении ярусов выделяют два (или три) яруса деревьев, один или два яруса кустарников, три яруса трав, один ярус напочвенного покрова. Всего получаем семь (или девять) ярусов.

Если считать, что ярусов семь, т.е. особо выделять три яруса трав и ярус напочвенного покрова, то имеем функцию

$$V(x,l,u,v,w) = \frac{k_0}{8}x^8 + lx^4 + ux^3 + vx^2 + wx.$$

[5] С точностью до написания черточек над x,l,u,v,w.

Это частный случай катастрофы типа «звезда» [9]:

$$V(x,p,q,l,u,v,w) = \frac{k_0}{8}x^8 + px^6 + qx^5 + lx^4 + ux^3 + vx^2 + wx. \tag{14}$$

Она описывает, вообще-то говоря, семь равновесных состояний сразу, из которых не более четырех являются устойчивыми (наличие сразу четырех локальных минимумов у функции V).

Литература

1. Москалюк, Т.А. Курс лекций по биогеоценологии [Электронный ресурс] / Т.А. Москалюк. – Режим доступа: http://www.botsad.ru/p_papers.htm (12.05.2009).
2. Грозовая активность как фактор пожарной опасности [Электронный ресурс] // Лесной эксперт. №4 (41), июнь 2007. – Режим доступа: http://www.lesnoyexpert.ru/
3. Кондратьев, К.Я. Лесные пожары как компонент природной экодинамики [Электронный ресурс] / К.Я. Кондратьев, Ал.А. Григорьев. – Режим доступа: http://www.nwicpc.ru/documents/Forest_fire-rus.DOC (12.05.2009).
4. Гуц, А.К. Математическая социология / А.К. Гуц, Л.А. Паутова, Ю.В. Фролова. – Омск:
5. Гуц, А.К. Математические методы в социологии / А.К. Гуц, Ю.В. Фролова. – М.: Издательство ЛКИ, 2007. 216 с.
6. Брёкер, Т. Дифференцируемые ростки и катастрофы / Т. Брёкер, Л. Ландер. – М.: Мир, 1977.
7. Постон, Т. Теория катастроф и ее приложения / Т. Постон, И. Стюарт. – М.: Мир, 1980.
8. Гилмор, Р. Прикладная теория катастроф / Р. Гилмор. – М.: Мир, 1984.
9. Woodcock, A.. A geometrical study of the elementary catastrophes / A. Woodcock, T. Poston // Lectures Notes in Math., N.373. – Springer, 1974.

Математические структуры и моделирование
2009, вып. 19, с. 68–77

УДК 519

КАТАСТРОФЫ ТИПА «ЛАСТОЧКИН ХВОСТ» В ЭКОЛОГИИ ЧЕЛОВЕКА

А.К. Гуц, Л.А. Володченкова

Предлагается качественное описание катастроф в экологии человека с помощью математической теории катастроф. Описываются состояния равновесия здоровья человека и их смена при изменении степени неблагоприятности/благоприятности внешних управляющих факторов.

Исследования по экологии человека предполагают наличие понятия здорового человека. Однако на сегодня нет общепринятой теории здоровья человека. П.И.Калью [1] к 1988 году рассмотрел 79 определений здоровья человека, сформулированные представителями различных научных дисциплин в разное время в различных странах мира. К 2009-му году их насчитsвается уже более ста.

В преамбуле Устава Всемирной организации здравоохранения (ВОЗ) индивидуальное здоровье определяется как «состояние полного физического, душевного и социального благополучия, а не только отсутствие болезней и физических дефектов».

Наличие самых различных формулировок понятия здоровья заставило классифицировать и представить как четыре модели здоровья человека:

1. «**Медицинская модель здоровья**. Она предполагает такое определение здоровья, которое содержит лишь медицинские признаки и характеристики здоровья. Здоровьем считают отсутствие болезней, их симптомов.

2. **Биомедицинская модель здоровья**. Здоровье рассматривается как отсутствие у человека органических нарушений и субъективных ощущений нездоровья. Внимание акцентируется на природно-биологической сущности человека, подчеркивается доминирующее значение биологических закономерностей в жизнедеятельности человека и в его здоровье.

3. **Биосоциальная модель здоровья**. В понятие здоровья включаются биологические и социальные признаки, которые рассматриваются в единстве, но при этом социальным признакам придается приоритетное значение.

4. **Ценностно-социальная модель здоровья**. Здоровье – ценность для человека, необходимая предпосылка для полноценной жизни, удовлетворения материальных и духовных потребностей, участия в труде и социальной жизни, в экономической, научной, культурной и других видах деятельности. Этой

Омский государственный университет.
E-mail: `guts@omsu.ru`, `volodchenkova@cmm.univer.omsk.su`

модели в наибольшей степени соответствует определение здоровья, сформулированное ВОЗ» [1] .

1. Описание модели здоровья человека

В рамках медицинской модели здоровья степень здоровья человека может быть охарактеризован достаточно большим числом количественных показателей, которые получают при проведении различных анализов (кровяное давление, температура тела, количество эритроцитов, сахар в крови и т.д.). К этим показателям следует добавить различные показатели, используемые другими моделями здоровья человека.

Путь величины x_j, $j = 1, 2, ..., N$ – совокупность всевозможных показателей здоровья человека.

Введем *интегральный показатель здоровья человека*, имеющий вид

$$x = \sum_{j=1}^{N} w_j x_j,$$

где w_j – вес показателя x_j, т.е. его вклад (доля) в интегральный показатель. Значения показателя x в момент времени t обозначаем как $x(t)$. Это число принимается нами как степень здоровья человека.

Показатель имеет нижнюю границу – число Z_0. Человек считается здоровым в момент времени t, если сумма его показателей $x(t) \geq Z_0$ и болеющим, если $x(t) < Z_0$.

Очевидно, что такой подход является крайне упрощенным, но любая модель здорового человека есть определенное упрощение, которое может быть со временем усложнена.

Здоровый в момент времени t человек без внешних (и ряда внутренних) причин в следующий момент времени $t + dt$ также будет зоровым.

Иначе говоря,

$$x(t + dt) = x(t) + A(t, x)dt, \tag{1}$$

где $A(t, x)dt$ – величина, описывающая отклонения в уровне здоровье человека, произошедшие на отрезке времени dt.

Из (1) получаем следующее дифференциальное уравнение:

$$\frac{dx}{dt} = A(t, x). \tag{2}$$

Очевидно, что в регионе, где проживают люди, может иметь место некоторый долговременный фактор риска для здоровья людей. Например, качество воды, или воздуха, или наличие радиоактивного фона.

В таком случае следует в правую часть дифференциального уравнения (2) добавить член $(-r)$:

$$\frac{dx}{dt} = A(t, x) - r. \tag{3}$$

В самом простом случае следует положить, что $A(t,x) = k_1 x$, т.е.

$$\frac{dx}{dt} = k_1 x - r. \tag{4}$$

Коэффициент k_1 можно посчитать постоянным. Но тогда степень здоровья человека будет нарастать как геометрическая прогрессия и явно «зашкалит», сделав бессмысленной нашу модель. Поэтому начнем ее усложнять.

Коэффициент k_1 отвечает за «прирост здоровья». Учитывая подходы биосоциальной и ценностно-социальной моделей здоровья человека, мы должны учесть, что здоровье человека зависит не только от его индивидуального здоровья, но от состояние здоровья его родителя(ей) и потомства[1]. Взаимодействие их здоровья есть величина $x \cdot x \cdot x$ – «произведение трех здоровий». Примем. что

$$k_1 = k_0 x \cdot x \cdot x - p(x) = k_0 x^3 - p(x), \tag{5}$$

где $k_0 = const$ – «сила» взаимодействия «трех здоровий», а величина $p(x)$ – это то, что мешает поддержанию всеобщего универсального зоровья человека-семьи. Это может быть один из факторов риска блогополучия здорового человека. Пусть это будет критический уровень k_2 медико-санитарной ситуации в регионе.

Тогда следует принять, что

$$p(x) = k_2 - k_3 x, \tag{6}$$

где член $k_3 x$ храктеризует принимаемые меры по преодолению неблагополучной медико-санитарной ситуации. То, что мы в этот член включили x, говорит о том, что принимаемые меры должны соотносится с состоянием здоровья людей в неблагополучном регионе.

Объединяя уравнения (4)-(6), получаем

$$\frac{dx}{dt} = k_0 x^4 - k_2 x + k_3 x^2 - r = \frac{\partial}{\partial x}\left\{\frac{k_0}{5}x^5 + \frac{k_3}{3}x^3 + \frac{-k_2}{2}x^2 + (-r)x\right\} \tag{7}$$

или

$$\frac{dx}{dt} = \frac{\partial}{\partial x} V(x,u,v,w), \tag{8}$$

где

$$V(x,u,v,w) = \frac{k_0}{5}x^5 + ux^3 + vx^2 + wx, \tag{9}$$

$$u = \frac{k_3}{3}, \quad v = -\frac{k_2}{2}, \quad w = -r.$$

Отметим, что хорошее здоровья людей характеризуется неравенством $x > Z_0$, ухудшение неравенством $x < Z_0$; действие долговременного вредоносного фактора риска неравенством $w < 0$, наличие неблагоприятной медико-санитарной

[1] Любой человек переживает болезнь своих родителей и детей, и это переживание сказывается на его собственном здоровье.

ситуации в регионе – неравенством $v < 0$, принятие мер по преодолению неблагополучной медико-санитарной ситуации (лечение) – неравенством $u > 0$.

Функция $V(x, u, v, w)$, заданная выражением (9), описывает при изменении параметров u, v, w самые различные бифуркации, называемые в математической теории катастроф катастрофами типа «ласточкин хвост» [2–5].

2. Равновесия системы, их смена и математическая теория катастроф

Равновесные системы, и это относится и экологии человека, характеризуются тем, что их строение и состав колеблются около какой-то средней точки, представляющей как бы типичное состояние состояния (здоровья) человека.

Равновесие в состоянии здоровья человека зависит от окружающей среды, а среда либо изначально неблагоприяная, либо постоянно подвержена изменениям. Неблагополучная медико-санитарная ситуация, антропогенные загрязнения, природные аномалии и т.д. оказывают влияние на организм человека. Человек либо заболевает, либо этого удается избежать, благодаря принятым мерам, и следовательно, сохранить состояние. В случае развития болезни вполне возможен переход к новому равновесию, при котором нельзя считать человека здоровым, но в болезненном состоянии он находится длительное время, осуществляя вместе с другми столь же больными людьми, общественное существование. И это действительно так: на Южном Урале имеется деревня, находящаяся с 1957 года в зоне сильного радиоактивного заражения, в которой люди живут, работают, создают семьи, рожают и выращивают детей, страдая и умирая при этом от раковых заболеваний. В Омске многие люди живут с 1970-х годов в бетонных домах, содержаших радиоактивный макеевский щебень. Известны также многочисленные примеры исторически длительного постоянного проживания людей в местности с плохой водой.

Радиактивный фон, плохая вода т.д. – это все внешние факторы, влияющие (и управляющиее) на экологию человека.

Для математического описания смены равновесий системы при изменении внешних (управляющих) факторов математиками создана теория элементарных катастроф. Она исходит из предположения, что поведение системы определяется некоторой потенциальной функцией $V = V(x, u_1, ..., u_n)$, где x – переменная, характеризующая изучаемую систему, а $u_1, ..., u_n$ – внешние управляющие факторы.

Динамика изменения во времени переменной $x(t)$ задается дифференциальным уравнением

$$\frac{dx}{dt} = \frac{\partial V(x, u_1, ..., u_n)}{\partial x} \tag{10}$$

Равновесия системы – это такие решения $x(t)$ уравнения (1), которые не меняются со сременем (какой-то период времени), т.е.

$$\frac{dx}{dt} = 0.$$

Но в таком случае, как видно из уравнения (1), имеем *уравнение всевозможных равновесий* системы
$$\frac{\partial V(x, u_1, ..., u_n)}{\partial x} = 0. \qquad (11)$$
Каждое равновесие есть решение x данного уравнения, зависящее от конкретного набора внешних управляющих факторов, т.е. $x = x(u_1, ..., u_n)$.

Если меняются внешние факторы, скажем вместо набора $(u_1^0, ..., u_n^0)$ переходим к набору $(u_1^1, ..., u_n^1)$, то имеем смену равновесия системы – вместо равновесия $x_0 = x(u_1, ..., u_n)$ получаем равновесие $x_1 = x(u_1^1, ..., u_n^1)$.

Естественно считать, что если слабо изменятся внешние факторы, то система почти не изменит своего состояния, т.е. практически сохранит значение характеризующей ее пременной x_0, или займет равновесие близкое к тому, что было, конечно же не отличающееся от предыдущего качественно. Однако исследования показали, что возможны такие малые изменения внешних факторов, и они происходят тогда, когда они пересекают при своем изменении так называемые *бифуркационные множества*, при которых система переходит к новому равновесию, для которого ее характерстика x_1 существенно отлична от предыдущей и, следовательно, новое равновесие обладает новыми качествами.

Такие переходы были названы катастрофами, поскольку переходы к новому *устойчивому* равновесию предшествует потеря устойчивости раннего равновесия. Название было продиктовано примерами из механики, физики и кораблестроения, где такие процессы описывали реальные катастрофы (перелом балки, замерзание воды, перевертывание судна). Отсюда и название математической теории, данное ей французским математиком Рене Томом.

С точки зрения математики, как видно из уравнения (11), равновесие $x = x(u_1, ..., u_n)$ – это либо *точка минимума*, либо *точка максимума*, либо так называемая *точка перегиба* функции $V = V_{(u_1,...,u_n)}(x) = V(x, u_1, ..., u_n)$. Обозначение $V_{(u_1,...,u_n)}(x)$ говорит о том, что мы функцию $V = V(x, u_1, ..., u_n)$ рассматриваем как функцию *одной* переменной x и используем соответствующий математический аппарат, известный школьникам.

Как правило, устойчивые равновесия системы – это минимумы функции $V = V_{(u_1,...,u_n)}(x)$. На рисунках графика функции $V = V_{(u_1,...,u_n)}(x)$ они изображаются ямками (соответственно максимумы – неустойчивые равновесия – изображаются вершинами горок).

Теория элементарных катастроф по самой своей природе локальна. Иначе говоря, функция $V_{(u,v,w)}(x) = V(x, u, v, w)$ как функция x определена лишь в окрестности $(-\varepsilon, +\varepsilon)$, и, следовательно, теория не описывает все возможные состояния равновесия. Уход системы в такие неописываемые теорией равновесия оговаривается как переход к равновесию $x = -\infty$ (минус бесконечность).

3. Определения экологический ситуации и экологических катастроф в экологии человека

Катастрофы, описываемые теорией катастроф вполне отвечают тому, что можно было бы назвать экологическими катастрофами. Поэтому адаптируем терминологию теории катастроф к экологии человека.

Назовем *местной экологической ситуацией*, отвечающей набору внешних факторов $(u_1, ..., u_n)$ любую из точек локального минимума функции $V_{(u_1,...,u_n)}(x)$.

Экосистема «здоровье человека» в случае местной экологической ситуации пребывает в *равновесии*, т.е. величина $x(t)$ не меняется со временем (управляющие факторы $(u_1, ..., u_n)$ зафиксированы).

Региональной экологической ситуацией, связанной с V, называется набор местных экологическая ситуация, предполагая и возможность равновесия $x = -\infty$ (минус бесконечность).

Принятие значение $x = -\infty$ означает, что наша математическая модель говорит о том, что экосистема «здоровье человека» переходит в равновесное состояние, которое не описывается с помощью предложенного уравнения ее эволюции (1) (теория локальна!).

Фактически региональная экологическая ситуация – это набор местных экологических ситуаций. Иначе говоря, это набор возможных равновесных состояний экосистемы «здоровье человека».

Точкой экологической катастрофы называется любая точка перегиба функции $V = V_{(u_1,...,u_n)}(x)$.

Морфологией катастрофы или множеством катастроф называется множество всех точек катастрофы.

Точки катастроф, которые входят в бифуркационное множество, находят, исключая x, в процессе решения системы уравнений

$$\frac{\partial V(x, u_1, ..., u_n)}{\partial x} = 0, \quad \frac{\partial^2 V(x, u_1, ..., u_n)}{\partial x^2} = 0.$$

Для конкретного набора внешних факторов у функции $V = V_{(u_1,...,u_n)}(x)$ может быть несколько точек минимумов. Скажем это точки $x_1, x_2, ..., x_m$. В каком из этих равновесий находится система? Для этого придуманы различные *правила*.

Правилом называется способ, по которому мы выбираем равновесия, т.е. минимумы функции $V = V_{(u_1,...,u_n)}(x)$ для некоторой региональной экологической ситуации, связанной с V.

Рассмотрим два основных правила.

Правило максимального промедления предписывает состоянию оставаться в минимуме при заменении факторов до тех пор, пока он существует (рис.1, вверху).

В момент, когда происходит исчезновение старого миниума и становится необходимым переход в новый, происходит экологическая катастрофа, т.е. резкое, скчкообразное изменение равновесия экосистемы «экология человека».

Правило Максвелла предписывает взять в качестве равновесия системы такую точку минимума, в которой функция $V = V_{(u_1,...,u_n)}(x)$ достигает наименьшего значения (рис.1, внизу).

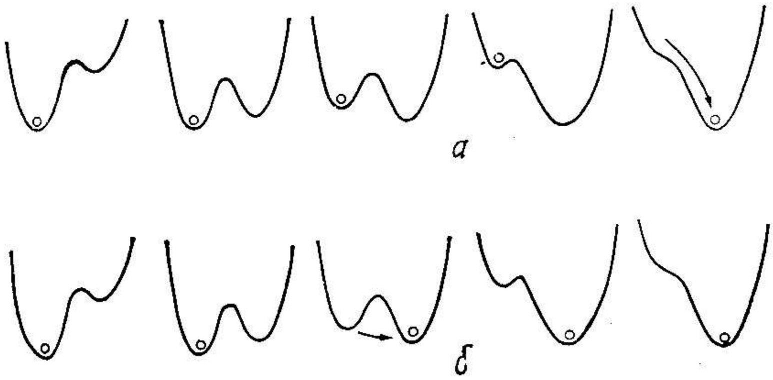


Рис. 1. Вверху показано действие правила максимального промедления, а на нижних рисунках – правило Максвелла.

Поскольку одним из значений этого минимума может оказаться $-\infty$, правилом Максвелла лучше пользоваться тогда, когда V имеет конечные минимумы. Ясно, что катастрофы возникают тогда, когда $V = V_{(u_1,\ldots,u_n)}(x)$ достигает абсолютного минимума в двух различных местах.

В случае правила максимального промедления катастрофе, т.е. резкой смене равновесия отвечает качественное изменение формы потенциальной функции $V = V_{(u_1,\ldots,u_n)}(x)$. В случае правила Максвелла резкая смена равновесия происходит без качественного изменения формы функции $V = V_{(u_1,\ldots,u_n)}(x)$. С точки зрения теории катастроф в случае правила Максвелла можно говорить о некатастрофической смене равновесия. Хотя с точки зрения экологии это скорее всего самая настоящая катастрофа. Катастрофы происходят в любом смысле и вне зависимости от принятого правила, если, изменясь управляющие внешние факторы пересекают бифуркционное множество.

Изучим все сказанное в нашем конкретном случае, когда функция V имеет вид (8).

4. Экологические катастрофы типа ласточкин хвост

Для катастрофы «ласточкин хвост» потенциал

$$V(x, u, v) = x^5 + ux^3 + vx^2 + wx.$$

Рассмотрим множество

$$M_V = \{(x, u, v, w) : \nabla V = 5x^4 + 3ux^2 + 2vx + w = 0\}.$$

Оно состоит из максимумом, минимумов и точек перегиба функции $V_{(u,v,w)}(x)$. Все они отвечают равновесию в состоянии изучаемой экосистемы человека.

Множество

$$S_V = \{(x, u, v, w) \in M_V : d^2V = 20x^3 + 6ux + 2v = 0\}$$

состоит из точек перегиба, которые принимаются, когда точка (u, v, w) принадлежит так называемому *бифуркационному множеству*:

$$B_V = \{(u, v, w) \in prS_V : \exists x(5x^4 + 3ux^2 + 2vx + w = 0 \ \& \ 20x^3 + 6ux + 2v = 0\}.$$

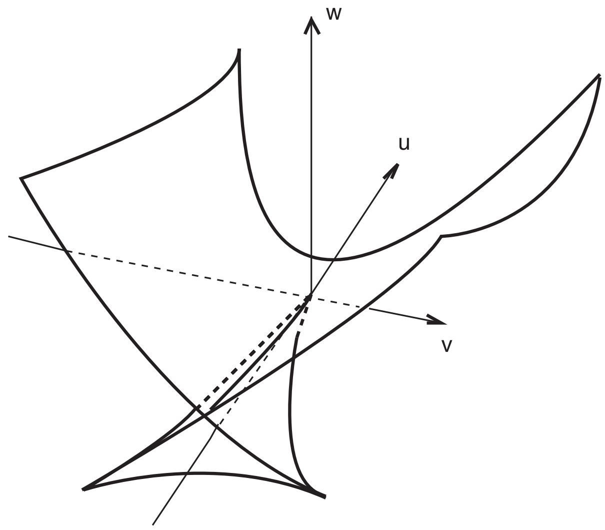


Рис. 2. Бифуркационное множество B_V для катастрофы «ласточкин хвост».

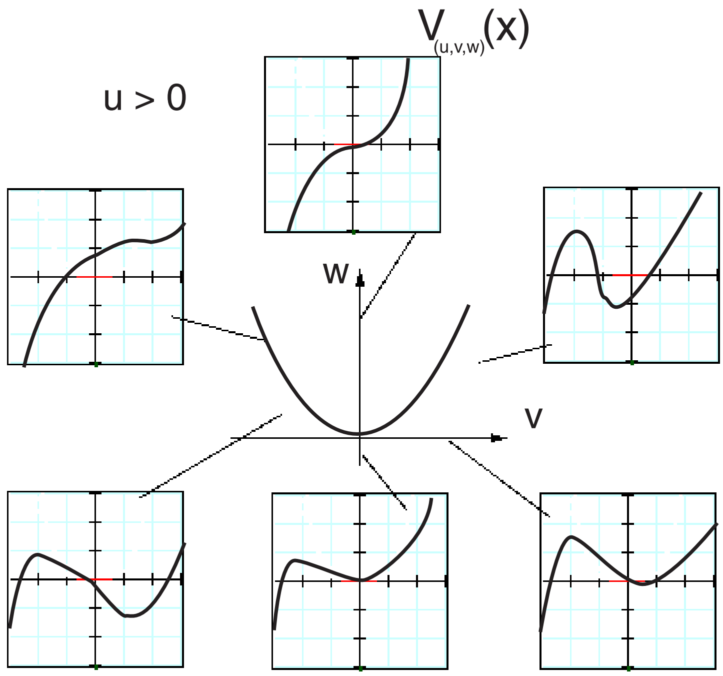


Рис. 3. Катастрофа «ласточкин хвост». Динамика появления и исчезновения максимумов и минимумов у функции $V_{(u,v,w)}(x)$ в сечении $u = const > 0$. При переходе через B_V рождается или умирает равновесие.

Множество B_V – это проекция множества S_V на плоскость (u, v, w). Оно изображено на (рис.2)

Если точка (u, v, w) меняется, т.е. меняются условия существования экосистемы человека, то сменяются состояния равновесия экосистемы. Изменения качественное, скачкообразное, катастрофическая если точка (u, v, w) при своем изменении перескает бифуркационное множество. Это и есть резкая смена состояния здоровья человека.

Для фиксированных управляющих факторов (параметров) (u, v, w) график функции $V_{(u,v,w)}(x)$ изображен на рис.3,4.

На рис.3 значение параметр $u > 0$ зафиксировано, а остальные два могут меняться. Бифуркционное множество на плоскости vw сводится к параболе. Раз $u > 0$, то принимаются меры по преодолению возможной неблагополучной медика-санитарной ситуации.

При этом, однако, даже в случае неблагополучной медика-санитарной ситуации ($v < 0$) и при действии долговременного вредоносного фактора риска неравенством $w < 0$ возможна устойчивая равновесная здоровая экологическая ситуация, т.е. $x > 0$ (точка (v, w) расположена ниже параболы, слева). В случае перехода точки (v, w) во внутрь параболу экологическая ситуации резко меняется, имеем экологическаую катастрофу, x уходит к $-\infty$, т.е. описание грядущего состояния экосистемы выходит за рамки нашей теории.

Так или иначе, но пользу лечения эта модель демонстрирует.

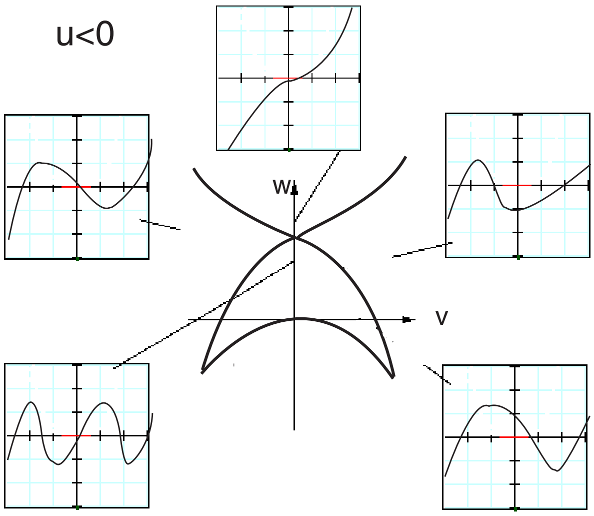


Рис. 4. Катастрофа «ласточкин хвост». Динамика появления и исчезновения максимумов и минимумов у функции $V_{(u,v,w)}(x)$ в сечении $u = const < 0$. При переходе через B_V рождается (умирает) одно или два равновесия.

Рассмотрим теперь случай $u < 0$, когда не принимаются меры по преодолению возможной неблагополучной медика-санитарной ситуации (см. 4). Мы видим, что даже в случае наличия неблагополучной медика-санитарной ситуации $v < 0$ возможно здоровое, т.е. $x > 0$, равновесное состояние экосистемы как при действии долговременного вредоносного фактора риска (график для V слева, с $x > 0$ в качестве локального минимума), так и при его отсутствии.

Но что более интересно: возможно также равновесное нездоровое состояние экосистемы при отсутствии неблагополучной медико-санитарной ситуации, т.е при $v > 0$ (график для V справа, с $x < 0$ в качестве локального минимума).

Наконец, возможны два альтернативных состояния равновесия – нездоровое и зоровое при отсутствии долговременного вредоносного фактора риска и при отсутствии/наличии неблагополучной медико-санитарной ситуации. Как тут не вспомнить южно-уральскую деревню.

5. Структурная устойчивость модели здоровья человека

Описание экологии человека, предложенное в данной статье, обладает важным достоинством – модель структурно устойчива. Иначе говоря, это означает, что найденная нами потенциальная функция (9)

$$V(x,u,v,w) = \frac{k_0}{5}x^5 + ux^2 + vx + w, \tag{12}$$

в случае ее возмущения (малого шевеления) вида

$$V(x,u,v,w) \to V(x,u,v,w) + \delta V(x,u,v,w)$$

может быть возвращена к выражению внешне такому же[2], как (12), если совершить следующие преобразования координаты x и параметров u, v, w:

$$x \to \overline{x} = \overline{x}(x,u,v,w),$$

$$(u,v,w) \to (\overline{u},\overline{v},\overline{w}) = \phi(u,v,w)$$

и прибавить к возмущенному потенциалу некоторую функцию $\psi(u,v,w)$ [4–6].

Иначе говоря, наша модель устойчива по отношению к изменениям потенциальной функции, которые могут проявляться в форме пожеланий внести уточнения в вид потенциальной функции, или в форме высказываний о том, что в реальности экология человека не может быть описан некоторой выбранной раз и навсегда конкретной математической формулой.

Наличие структурной устойчивости у модели говорит о том, что ничего качественно нового в описании экологических ситуаций и экологических катастроф возмущенная модель не дает.

Литература

1. Калью, П.И. Сущностная характеристика понятия «здоровье» и некоторые вопросы перестройки здравоохранения: обзорная информация / П.И. Калью. – М., 1988. 220 с.
2. Гуц, А.К. Математическая социология / А.К. Гуц, Л.А. Паутова, Ю.В. Фролова. – Омск:
3. Гуц, А.К. Математические методы в социологии / А.К. Гуц, Ю.В. Фролова. – М.: Издательство ЛКИ, 2007. 216 с.
4. Брёкер, Т. Дифференцируемые ростки и катастрофы / Т. Брёкер, Л. Ландер. – М.: Мир, 1977.
5. Постон, Т. Теория катастроф и ее приложения / Т. Постон, И. Стюарт. – М.: Мир, 1980.
6. Гилмор, Р. Прикладная теория катастроф / Р. Гилмор. – М.: Мир, 1984.

[2] С точностью до написания черточек над x, u, v, w.

Математические структуры и моделирование
2009, вып. 19, с. 78–84

УДК 530.12

СПОСОБ РАСЧЕТА СКАЧКОВ ЭНЕРГИИ ПРИ ИЗМЕНЕНИИ РАЗМЕРНОСТИ ПРОСТРАНСТВА И ВРЕМЕНИ

А.К. Гуц

Предлагается способ расчёта скачков плотности энергии при изменении размерности пространства и времени (соответственно, при изменении размерности пространства-времени).

Мы ощущаем, а физические эксперименты это подтверждают, что размерность нашего физического пространства равна 3. Это базовая размерность, и 3-мерное пространство является базовым физическим пространством, к которому привязано человеческий способ осознания (и созидания) Вселенной.

То же мы можем сказать о времени. Мы уверены, что оно одномерно. Это базовая размерность времени.

Следовательно 4-мерное лоренцево многообразие – 4-мерное пространство-время – это базовая модель окружающей нас Реальности, часть которой являемся и мы сами.

Однако вполне допустимы локальные и глобальные изменения размерности физического пространства или времени. Размерность – это топологическая характеристика пространства. С ней связан конкретный набор чисел Бетти. Для 3-мерного пространств 3-мерное Бетти не равно нулю, а все m-мерные числа Бетти β_m с $m \geq 4$ являются нулевыми. Если вдруг 4-мерное число Бетти окажется отличным от нуля, то это говорит о том, что физическое пространство стало 4-мерным.

Ясно, что с точки зрения физики спонтанные (или искусственные) изменения размерности пространства или времени должны характеризоваться скачками (плотности) энергии.

Как оценить эти скачки плотности энергии?

Существуют различные интегральные формулы (интегралы по многообразию), связывающие скаляры, образованные из скалярной кривизны, секционных кривизн, сверток тензора Риччи и тензора кривизны с самими собой, с числами Бетти, с характеристикой Эйлера-Пуанкаре, с числами Понтрягина, с сигнатурой (формула Хинценбруха). Такие формулы называют чаще всего

Омский государственный университет.
E-mail: `aguts@mail.ru`

формулами Черна-Гаусса-Бонне. Подынтегральные выражения в этих формулах можно выразить через плотности энергии, которая входит в уравнения гравитационного поля, принимаемой к рассмотрению теории гравитации. Этой теорией может быть традиционная теория гравитация Эйнштейна или, к примеру, теория гравитация Эйнштейна-Гаусса-Бонне.

Пишем одну формулу Черна-Гаусса-Бонне для базового пространства, а вторую – для пространства с измененной размерностью. Обе дают оценку усредненной плотности энергии; одна для базового пространства, а вторая для изменённого. Разница и есть искомый скачок энергии, характеризующий изменение такой топологической характеристики пространства-времени, как размерность.

Ясно, что существенные скачки плотности энергии могут быть обнаружены в астрономических наблюдениях обширных областей физического пространства. Поэтому не исключено, что размерность пространства-времени в таких областях Вселенной может быть отличной от 4, и а размерность физического пространства вообще меняться с течением времени.

1. Формула Черна-Гаусса-Бонне для псевдоримановых многообразий M^{2k}

Пусть $< M^{2k}, g >$ – $2k$-мерное компактное ориентированное псевдориманово многообразие сигнатуры $< \underbrace{+...+}_{p}\underbrace{-...-}_{2k-p} >$.

Пусть $dv = \sqrt{|det(g)|}dx^1 \wedge ... \wedge dx^{2k}$ – $2k$-форма объема и

$$P_k(R) = \frac{(-1)^{[p/2]}}{2^{2k}(2\pi)^k k!}\varepsilon^{i_1 i_2 ... i_{2k}}_{j_1 j_2 ... j_{2k}} Ri_1i_2j_1j_2 Ri_3i_4j_3j_4 ... Ri_{2k-1}i_{2k}j_{2k-1}j_{2k}$$

где

$$\varepsilon^{i_1 i_2 ... i_{2k}}_{j_1 j_2 ... j_{2k}} = \begin{cases} +1, \{i_1 i_2 ... i_{2k}\} & \text{четная перестановка,} \quad \{j_1 j_2 ... j_{2k}\} \\ -1, \{i_1 i_2 ... i_{2k}\} & \text{нечетная перестановка,} \quad \{j_1 j_2 ... j_{2k}\} \\ \ \ 0, \quad \text{среди} \quad \{i_1 i_2 ... i_{2k}\} & \text{или среди} \{j_1 j_2 ... j_{2k}\} \quad \text{есть одинаковые} \end{cases}$$

Тогда $[1, 2]$

$$\int\limits_{M^{2k}} P_k(R)dv = \chi(M^{2k}). \tag{1}$$

2. Скачки размерности пространства-времени. Случай замкнутого мноообразия

Из уравнений поля для $2k$-мерного псевдориманова пространства, т.е. для $2k$-мерного пространства-времени

$$R^{(2k)}_{ik} - \frac{1}{2}g^{(2k)}_{ik}R^{(2k)} = \kappa\varepsilon_{(2k)}u_i u_k,$$

а также из структуры формулы для компонент тензора кривизны, получаем, что
$$R^{(2k)} \sim \kappa\varepsilon_{(2k)}, \quad R^{(2k)}_{ik}[R^{(2k)}]^{ik} \sim \kappa\varepsilon_{(2k)}, \quad R^{(2k)}_{iklm}[R^{(2k)}]^{iklm} \sim \kappa\varepsilon_{(2k)}.$$
Поэтому
$$Pf_k(R) \sim (\kappa\varepsilon_{(2k)})^k.$$
Следовательно, с помощью формулы Гаусса-Бонне-Черна (1) получаем следующую оценку для среднего значения плотности энергии $2k$-мерной реальности:
$$(\kappa\langle\varepsilon_{(2k)}\rangle)^k vol(M^{2k}) \sim \chi(M^{2k})$$
Предположим, что базовым является 4-мерное пространство-время M^4, для которого $k = 2$, и объем $vol(M^4) \sim l^4$.

Предположим также, что Реальность такова, что размерность пространства-времени как модели Реальности «колеблется спонтанно» около размерности базового пространства-времени, т.е. следует наравне с M^4 рассматривать модели M^{2k}. При этом топология M^{2k} может иметь вид $M^4 \times S^1 \times ... \times S^1$. Примем, что дополнительные размерности характеризуются величиной λ, малой по сравнению с числом l, т.е. $l/\lambda > 1$, и совершенно не важно идет ли речь о пространственном дополнительном измерении или о временном.

Таким образом, при $k \geq 2$
$$vol(M^{2k}) \sim l^4\lambda^{2k-4} = l^{2k}(\lambda/l)^{2k-4}$$
и поэтому
$$\langle\varepsilon_{(2k)}\rangle \sim \frac{1}{\kappa l^2}[\chi(M^{2k})]^{1/k}(l/\lambda)^{2-4/k}.$$
Следовательно, так как $l/\lambda > 1$, то
$$\langle\varepsilon_{(4)}\rangle \leq \langle\varepsilon_{(6)}\rangle \leq ... \leq \langle\varepsilon_{(2k)}\rangle \leq ...$$
Таким образом, мы видим, что переход к более мерному пространству-времени требует «затрат» энергии, а уменьшение размерности означат «сброс» энергии. При этом 4-мерная Реальность, которую мы назвали базовой, является основным энергетическим уровнем с минимальной средней энергией.

3. Вероятности переходов при смене размерности

Как вычислить вероятности переходов $M^{2k} \to M^{2p}$? Для этого воспользуемся подходом, предложенным в книге [3].

Реальность может представляться нам как псевдориманово многообразие M^n любой размерности n в зависимости от нашего способа созидания Реальности и ее осознания.

Фактически мы считаем, что переходы $M^{2k} \to M^{2p}$, влекущие скачки размерности прстранства и времени, происходят не в силу того, что это некоторый естественный природный процесс, а в силу того, что люди вынуждают Реальность (природу) это совершать в силу скачкообоазной эволюции своих представлений о том как устроена эта Реальность. Смена представлений – это смена способа осознания Реальности, смена своих идей-фантазий о структуре Реальности [4].

Базовым является 4-мерное лоренцево многообразие $< R^4, g^{(4)} >$. Обозначаем способы осознания как $\ell A, \ell B, ...$. В [3] способы осознания формализуются как гладкие кольца:

$$\ell A = C^{\infty}(\mathbb{R}^m), \quad \ell A = C^{\infty}(\mathbb{R}^n), ...$$

В случае способа осознания ℓA Реальность предстает как $(4+m)$-мерная среда $R^4_{\ell A}$ с метрикой $g^{(4)}(\ell A)$:

$$\sum_{I,J=1}^{4+m} g^{(4)}(\ell A)_{IJ} dz^I dz^J = g^{(4)}_{ik}(x^0, ..., x^3, a) dx^i dx^k + 2 s_{il} dx^i da^l + h_{kl} da^k da^l, \qquad (2)$$

$$z = (x, a), \quad x \in \mathbb{R}^4, \; a \in \mathbb{R}^m.$$

Имеем для амплитуды вероятности перехода

$$\langle g^{(4)}_1(\ell A) | g^{(4)}_2(\ell B) \rangle = \int\limits_{g^{(4)}_1(\ell A)}^{g^{(4)}_2(\ell B)} \mathcal{D}[g^{(5)}(\ell A, \ell B)] e^{\frac{i}{\hbar} S[g^{(5)}(\ell A, \ell B)]}, \qquad (3)$$

где

$$S[g^{(5)}(\ell A, \ell B)] = \kappa_m \int \sqrt{|g^{(5)}(\ell A, \ell B)|} R^{(5)}(\ell A, \ell B) d^5 x. \qquad (4)$$

Для того чтобы написать эти формулы в более осмысленном виде необходимо учесть наличие морфизма $\Phi : \ell B \to \ell A$ между стадиями $\ell A = \ell C^{\infty}(\mathbb{R}^n)$ и $\ell B = \ell C^{\infty}(\mathbb{R}^m)$. Это означает, что $a = \phi(b)$, где $a \in \mathbb{R}^n, \; b \in \mathbb{R}^m$.

Иначе говоря, вместо (3)-(4) пишем

$$\langle g^{(4)}_1(\ell A) | g^{(4)}_2(\ell B) \rangle = \int\limits_{g^{(4)}_1(x, \phi(b))}^{g^{(4)}_2(x, b)} \mathcal{D}[g^{(5)}(b)] e^{\frac{i}{\hbar} S[g^{(5)}(b)]}, \qquad (5)$$

где

$$S[g^{(5)}] = \kappa_m \int \int \sqrt{|g^{(5)}(x, b)|} R^{(5)}(x, b) d^5 x d^m b. \qquad (6)$$

4. Формула Черна-Гаусса-Бонне для псевдоримановых многообразий M^{2k} с краем

Пусть M^{2k} псевдориманово мнгообразие с краем ∂M^{2k} и $q : T_1 M^{2k} \to M^{2k}$ расслоение на сферы, ассоциированное с касательным расслоением TM^{2k} (т.е.

состоящее из векторов касательного расслоения с нормой 1). Существует $(2k-1)$-форма σ на T_1M^{2k}, для которой

$$\int\limits_{q^{-1}(x)} \sigma = 1 \ \text{ для всех } x \in M^{2k}$$

и $q^*Pf_k(\Omega) = d\sigma$.

Всякое векторное поле T, нормальное к ∂M^{2k}, задает несингулярное сечение $\tau : \partial M^{2k} \to T_1M^{2k}$.

Тогда имеет место формула Черна-Гаусса-Бонне для псевдоримановых многообразий M^{2k} с краем [6]:

$$\int\limits_{M^{2k}} Pf(\Omega) = ind_{\partial M^{2k}}T + \int\limits_{\partial M^{2k}} \tau^*\sigma, \tag{7}$$

где $ind_{\partial M^{2k}}$ определяется следующим образом.

Если T не нулевое векторное поле на ∂M^{2k}, $\overline{T}$ – продолжение векторного поля T на все мнгообразие M^{2k} и $a_1, ..., a_k$ – конечное число особых точек (нулей) поля $\overline{T}$ на M^{2k} [5, c.516], то

$$ind_{\partial M^{2k}}T = \sum_{j=1}^{k} ind_{a_j}\overline{T}.$$

Если поле T трасверсально (в частности, нормально) к ∂M^{2k}, то

$$ind_{\partial M^{2k}}T = \chi(M^{2k}).$$

В общем случае, для ориентированного компактного многообразия с краем:

$$ind_{\partial M^{2k}}T = \chi(M^{2k}) - deg(K_T)$$

где $deg(K_T)$ – степень отображения $K_T : \partial M^{2k} \to S^{2k-1}$ [5, c.502], $K_T(x)$ равно точке на S^{2k-1}, отмеченной концом вектора $v = v^0T + v^1u_1 + ... + v^{2k-1}u_{2k-1}$, $\{u_1, ..., u_{2k-1}\}$ – базис в $T_x(\partial M^{2k})$, $x \in \partial M^{2k}$, $\sum_j v^j = 1$ [7].

Если Q^{2k} – $2k$-мерная компактная область в M^{2k} с границей ∂Q^{2k}, то формулу (7) можно переписать в виде

$$\int\limits_{Q^{2k}} Pf(\Omega) = ind_{\partial Q^{2k}}T + \int\limits_{\partial Q^{2k}} \tau^*\sigma, \tag{8}$$

где все формы, поля, определяются как выше с заменой буква M на букву Q.

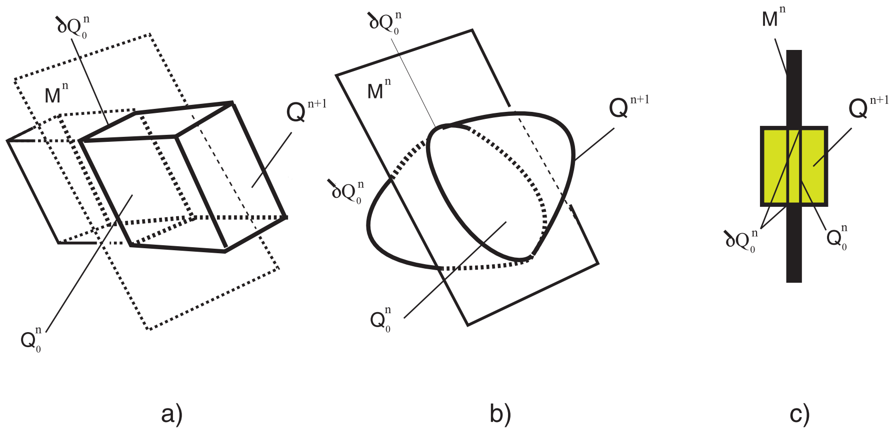


Рис. 1. a) Многообразие M^n, в котором клетка Q_0^n с границей (краем) ∂Q_0^n становится внутренней частью клетки Q^{n+1} большей размерностью; b) сглаженный вариант левого рисунка ; c) вид «сбоку» на процедуру увеличения размерности многообразия в «месте» Q_0^n.

5. Скачки размерности пространства и времени. Общий случай

Пусть у пространства-времени M^n размерности n область Q^n с границей (краем) ∂Q^n становится внутренней часть более мерной области Q^{n+1} (см. рис.1).

Поскольку формула Гаусса-Бонне-Черна нетривиальна для четномерных многообразий, то считаем, что $n = 2k$, и включаем область Q^{n+1}, как часть, в $(n+2)$-мерную область Q^{n+2}.

Теперь для того чтобы оценить скачки энергии нам нужна выписать формулы Гаусса-Бонне-Черна для мнообразия с краем (8) – одна для пары $(Q^n, \partial Q^n)$, а другая для пары $(Q^{n+2}, \partial Q^{n+2})$.

Следует заметить, что получение точных оценок скачков энергии на пути, указанном в данной заметке, является весьма трудной математической задачей, поскольку сложно выразить геометрические члены, входящие в формулы типа Черна-Гаусса-Бонне, через плотность энергии, входящую в уравнения поля.

6. Расчет изменения размерности физического пространства

Физическое прстранство моделируется математически как риманово многообразие, т.е. как пара $< V^n, g >$, где V^n – гладкое многообразие, а g – положительно определенная риманова метрика.

Мы можем считать, глядя на Реальность как на Нечто, находящееся в пространстве, что это Нечто меняется с течением времени t. Таков классический, доминковианский взгляд на Реальность.

Физическое пространство может иметь размерность $\dim = n$ с базовым значением $\dim = 3$. Предположим, что размерность в момент времени t_0 меняется и вместо значения n принимает значение m. Какая для этого требуется энергия?

Ясно, что расчет скачка энергии можно проивести по схеме, изложенной в предыдущих параграфах, с той только разницей, что пишутся формулы Гаусса-

Бонне-Черна для римановых многообразий $< V^n, g^{(n)} >$ и $< V^m, g^{(m)} >$ и используются пространственные компоненты уравнения поля.

Литература

1. Avez, A. Formula de Gauss-Bonnet-Chern en métrique de signature quelconque / A. Avez // C.R. Acad. Sci. Paris. – 1962. – T.255. – P.2049-2051.
2. Chern, S.S. Pseudo-Riemannian Geometry and the Gauss-Bonnet Formula / S.S. Chern // Ann. Acad. Brasil Ci. – 1963. – V.35. – P.17-26.
3. Гуц, А.К. Элементы теории времени / А.К. Гуц. – Омск: Издательство Наследие. Диалог-Сибирь, 2004. 364 с.
4. Гуц, А.К. Основы квантовой кибернетики: Учебное пособие / А.К. Гуц. – Омск: Полиграфический центр КАН, 2008. 204 с
5. Дубровин, Б.А. Современная геометрия / Б.А. Дубровин, С.П. Новиков, Ф.Т.Фоменко. – М.:Наука, 1979.
6. Pelletier, F. Quelques proprietes géométriques des variétés pseudo-riemanniennes singuliéres / F. Pelletier // Annales de la faculé des sciences de Toulouse 6^e série. – 1995. – T.4, N.1. – 87-199.
7. Alty, L.J. The Generalised Gauss-Bonnet-Chern Theorem / L.J. Alty // J.Math.Phys. – 1995. – V.36. – P.3094-3105.

Математические структуры и моделирование
2009, вып. 19, с. 85–96

УДК 65.012.8

ТЕОРЕТИКО-ГРАФОВЫЙ АНАЛИЗ РОЛЕВОЙ ПОЛИТИКИ БЕЗОПАСНОСТИ

С.В. Белим, С.Ю. Белим, Н.Ф. Богаченко

В работе рассматривается формальная модель ролевого разграничения доступа. Представлен классовый подход к построению иерархии ролей, являющийся обобщением классического листового принципа распределения полномочий в ролевой политике безопасности. Показана возможность построения данной политики на произвольном ориентированном графе ролей.

Проблема разграничения доступа к данным является ключевым элементом систем безопасности компьютерной информации. Анализ теоретических исследований и программно-технических разработок в области информационной безопасности позволяет выделить четыре основных подхода в построении политики разграничения доступа [1]: дискреционные, мандатные, тематические и ролевые модели.

Особый интерес представляет ролевая политика безопасности. Это обусловлено тем, что на практике получила широкое распространение технология рабочих групп пользователей в системах разграничения доступа, являющаяся упрощенным вариантом ролевой политики, тогда как формальная ролевая модель изучена не достаточно полно.

1. Ролевая политика безопасности

Рассмотрим базовую модель ролевого разграничения доступа [1, 2], характеризующуюся неизменными установками системы безопасности. Выделим основные элементы, которые понадобятся для дальнейшего описания.

Основные множества:

1. U – множество пользователей системы.
2. R – множество ролей, возможных в системе.
3. P – множество полномочий (прав) на действия в системе.

Омский государственный университет.
E-mail: nfbogachenko@mail.ru

Зададим *отображения*, определяющие функционирование системы под управлением ролевой политики безопасности:

1. $RP : R \to 2^P$ – множество полномочий для роли, при этом $\forall\ p \in P\ \exists\ r \in R :\ p \in RP(r)$.

2. $UR : U \to 2^R$ – множество ролей, на которые может быть авторизирован пользователь. Следует отметить, что возможно существование ролей, на которые не авторизирован ни один пользователь.

Определение 1. ***Иерархией ролей*** будем называть отношение частичного нестрогого порядка, заданное на множестве ролей R. При этом, если $r_2 \leq r_1$, то r_1 находится в иерархии ролей «выше», чем r_2 [1].

Исходя из определения, иерархию ролей можно представить в виде ориентированного графа $G = (R, E)$. Множество вершин R – это множество ролей. Дуга $(r_1, r_2) \in E$, если в иерархии ролей $r_2 \leq r_1$ [2].

В большинстве работ, посвященных ролевому разграничению доступа, принято считать, что иерархия ролей имеет вид ориентированного дерева $[1, 2]$, в дальнейшем будем называть его ***деревом ролей*** и обозначать $T = (R, E)$.

При иерархическом отношении ролей важным является вопрос построения отображения RP, задающего распределение множества полномочий для ролей, а именно, возможно ли назначение одного и того же набора полномочий двум ролям, находящимся в иерархическом подчинении. При этом применяется механизм наследования «снизу – вверх»: назначение полномочий начинается с листовых вершин – ***листовое распределение прав доступа***. Для этого случая возможны три подхода к построению отображения RP [1]:

1. Строго таксономический листовой подход. Все множество полномочий разбивается на непересекающиеся подмножества:

$$P = \bigcup_{i=1}^{k} P_i,\ \forall i, j \in \{1, \ldots, k\} :\ P_i \cap P_j = \varnothing. \tag{1}$$

Здесь k – количество листовых вершин дерева ролей. Пусть R_L ($R_L \subset R$) – множество листовых вершин. Каждой листовой вершине дерева ролей r_i отображение RP сопоставляет одно из подмножеств P_i:

$$\forall r_i \in R_L : RP(r_i) = P_i. \tag{2}$$

[1]Необходимо также выполнение еще одного условия: $\forall u \in U : (r, r' \in R) \wedge (r \in UR(u)) \wedge (r' \leq r) \Rightarrow (r' \in UR(u))$. То есть, вместе с заданной ролью пользователь должен быть авторизирован и на все роли, которые лежат в иерархии ниже.

[2]Если следовать системе обозначений графического языка моделирования UML (Unified Modeling Language), позволяющего представлять различные объектно-ориентированные проекты в единых обозначениях, то дугу надо ориентировать от младшей роли r' – к старшей r. Чтобы сохранить терминологию теории графов, будем считать, что дуги направлены от старших ролей – к младшим. Очевидно, в обоих случаях отношение порядка, задающее иерархию ролей, и соответствующий неориентированный граф останутся неизменными.

Полномочия остальных (нелистовых) вершин дерева ролей определяются как объединение полномочий непосредственно подчиненных им ролей:

$$\forall r \notin R_L : RP(r) = \bigcup_{r' \in Ch(r)} RP(r'), \tag{3}$$

где $Ch(r)$ – полный набор сыновей вершины r.

2. Нетаксономический листовой подход. Полномочия непосредственно получают также только листовые роли, а полномочия остальных ролей получаются объединением полномочий сыновей. Однако, допускается непустое пересечение полномочий листовых ролей:

$$P = \bigcup_{i=1}^{k} P_i,\ \exists i, j \in \{1, \ldots, k\} :\ P_i \cap P_j \neq \varnothing. \tag{4}$$

3. Иерархический охватный листовой подход. В набор полномочий старших ролей непосредственно включаются (добавляются) только те полномочия, которые не вошли в наборы полномочий подчиненных ролей.

За счет иерархической структуры, во всех трех случаях в итоговом наборе полномочий присутствуют все полномочия подчиненных ролей.

2. Разбиение ролей на классы эквивалентности

Пусть иерархия ролей задана в виде ориентированного дерева $T = (R, E)$. Определим разбиение множества листовых вершин R_L дерева ролей T на k подмножеств:

$$R_L = \bigcup_{i=1}^{k} R_L^{(i)},\ \forall i, j \in \{1, \ldots, k\} : R_L^{(i)} \cap R_L^{(j)} = \varnothing. \tag{5}$$

Данное разбиение задает отношение эквивалентности $\overset{R_L}{\backsim}$ на множестве листовых вершин.

Рассмотрим теперь ***классовое распределение прав доступа***. Пусть две роли, относящиеся к одному классу эквивалентности листовых вершин, имеют одинаковые права:

$$\forall r_1, r_2 \in R_L : (r_1 \overset{R_L}{\backsim} r_2) \Rightarrow (RP(r_1) = RP(r_2)). \tag{6}$$

Как и для традиционного листового распределения полномочий, возможны три алгоритма постоения отображения RP на всем множестве ролей R.

1. Строго таксономический классовый подход. Разобъем множество P на k непересекающихся подмножеств по числу классов эквивалентности листовых вершин дерева ролей:

$$P = \bigcup_{i=1}^{k} P_i,\ \forall i, j \in \{1, \ldots, k\} :\ P_i \cap P_j = \varnothing. \tag{7}$$

Распределение прав, определяющееся отображением $RP : R \to 2^P$, зададим для листовых вершин в следующем виде:

$$\forall r \in R_L^{(i)} : RP(r) = P_i. \tag{8}$$

Для нелистовых вершин множество прав будем определять как объединение прав всех вершин, которые являются сыновьями данной вершины:

$$\forall r \notin R_L : RP(r) = \bigcup_{r' \in Ch(r)} RP(r'). \tag{9}$$

Здесь, как и прежде, через $Ch(r)$ обозначено множество всех сыновей вершины r.

2. Нетаксономический классовый подход. Аналогично предыдущему случаю, рапределение прав изначально производится только между листовыми вершинами, но множества прав различных классов эквивалентности листовых вершин могут пересекаться.

3. Иерархический охватный классовый подход. Распределение прав производится между классами эквивалентности листовых вершин и передается по иерархическим принципам. Но, кроме того, нелистовые вершины, унаследовавшие одинаковые наборы прав, могут одновременно получать дополнительные права.

Очевидно, что каждый из трех листовых подходов распределения полномочий является частным случаем соответствующего классового подхода при условии: $|R_L^{(i)}| = 1$, где $\{R_L^{(i)}\}_{i=1}^k$ – начальное разбиение множества листовых вершин, другими словами, каждая листовая вершина образует отдельный класс разбиения.

Оказывается, что разбиение листовых вершин дерева ролей вместе с правилами построения отображения RP порождает разбиение всего множества ролей.

Определение 2. Будем считать две роли ***эквивалентными***, если они наделены одинаковыми правами:

$$\forall r_1, r_2 \in R : (RP(r_1) = RP(r_2)) \Rightarrow (r_1 \overset{RP}{\backsim} r_2). \tag{10}$$

Определение 3. Отношение эквивалентности $\overset{RP}{\backsim}$ задает разбиение множества ролей R. Полученные классы эквивалентности ролей будем называть RP***-классами***.

Определение 4. Каждому узлу r дерева ролей припишем соответствующий данной роли набор полномочий $RP(r)$. Результирующее помеченное дерево ролей назовем RP***-деревом***.

Итак, в RP-дереве в один RP-класс попадают вершины, помеченные одним и тем же набором полномочий.

Обозначим число RP-классов через K, а число классов эквивалентности листовых вершин через k. Очевидно, что в случае строгого таксономического классового подхода $K \geqslant k$. При двух других подходах возможно получить $K < k$ в тех ситуациях, когда несколько классов эквивалентности листовых вершин наделяются одним и тем же набором полномочий.

3. Оптимальные RP-деревья

Определение 5. RP-дерево называется ***вырожденным***, если на нем существует ровно один RP-класс: $K = 1$.

Например, вырожденным является RP-дерево, представляющее собой ориентированную цепь, в случае строго таксономического классового подхода.

Определение 6. RP-дерево называется ***оптимальным***, если количество заданных на нем RP-классов совпадает с количеством вершин: $K = |R|$.

Заметим, что неоптимальность может быть присуща не только RP-деревьям, полученным в результате классового распределения полномочий. Она может появиться и при листовом построении отображения RP.

Очевидно, что для оптимальности RP-дерева необходимо потребовать, чтобы в начальном разбиении множества листовых вершин каждый лист составлял отдельный класс (использовалось листовое распределение полномочий). Далее рассмотрим необходимые и достаточные условия оптимальности.

Теорема 1. *При строго таксономическом листовом подходе распределения прав, RP-дерево является оптимальным тогда и только тогда, когда полустепень исхода (число исходящих дуг) каждой нелистовой вершины не меньше двух:*

$$\forall r \notin R_L : d^-(r) \geqslant 2. \tag{11}$$

Доказательство. Пусть RP-дерево является оптимальным. Тогда

$$\forall r_1, r_2 \in R : (r_1 \neq r_2) \Rightarrow (RP(r_1) \neq RP(r_2)). \tag{12}$$

От противного. Пусть $\exists r_1 \notin R_L : d^-(r_1) < 2$, следовательно $d^-(r_1) = 1$. Тогда вершина r_1 имеет ровно одного сына, обозначим его r_2. Согласно правилам строго таксономического листового подхода: $RP(r_1) = RP(r_2)$ – противоречие.

Пусть теперь выполнено неравенство (11). Рассмотрим две различные вершины $r_1, r_2 \in R$. Надо показать справедливость условия (12). Заметим, что требование (11) влечет выполнение следующего неравенства:

$$\forall r_1, r_2 \in R : (r_1 \neq r_2) \Rightarrow (R_L(r_1) \neq R_L(r_2)), \tag{13}$$

где $R_L(r_i)$ – потомки вершины r_i, являющиеся листовыми вершинами в случае, когда $r_i \notin R_L$, либо сама вершина r_i, если она листовая. Данное утверждение очевидным образом следует из ацикличности принятой иерархии ролей (T – дерево). Согласно введенной системе обозначений, при строго таксономическом листовом подходе:

$$\forall r \in R : RP(r) = \bigcup_{r' \in R_L(r)} RP(r') \tag{14}$$

и

$$\forall r', r'' \in R_L : (r' \neq r'') \Rightarrow (RP(r') \cap RP(r'') = \varnothing). \tag{15}$$

Из (13) и (15) следует:

$$\forall r_1, r_2 \in R : (r_1 \neq r_2) \Rightarrow \left(\bigcup_{r' \in R_L(r_1)} RP(r') \neq \bigcup_{r' \in R_L(r_2)} RP(r') \right). \qquad (16)$$

Принимая во внимание равенство (14), получаем условие (12). ■

Теорема 2. *При нетаксономическом листовом подходе распределения прав, RP-дерево является оптимальным тогда и только тогда, когда полустепень исхода каждой нелистовой вершины не меньше двух (выполнено условие (11)) и разбиение множества прав P на подмножества P_i произведено таким образом, что*

$$\forall j \in \{1, \ldots, k\} : \ P_j \nsubseteq \bigcup_{i=1, i \neq j}^{k} P_i. \qquad (17)$$

Доказательство. Необходимость доказывается аналогично предыдущей теореме. При доказательстве достаточности условие (15) заменяется условием (17). Пусть $I_1, I_2 \subseteq \{1, \ldots, k\} : (I_1 \neq I_2)$, тогда $\exists j : (j \in I_1) \wedge (j \notin I_2)$. Согласно (17): $P_j \nsubseteq \bigcup_{i \in I_2} P_i$, следовательно $\bigcup_{i \in I_1} P_i \nsubseteq \bigcup_{i \in I_2} P_i$. В результате:

$$\forall I_1, I_2 \subseteq \{1, \ldots, k\} : (I_1 \neq I_2) \Rightarrow \left(\bigcup_{i \in I_1} P_i \neq \bigcup_{i \in I_2} P_i \right). \qquad (18)$$

Тогда, принимая во внимание неравенство (13) и то, что $\forall r_i \in R_L : RP(r_i) = P_i$, получаем (16) и, как следствие, (12). ■

Отметим, что при иерархическом охватном листовом подходе оптимальным может быть RP-дерево произвольной структуры (например, ориентированная цепь) за счет того, что нелистовые вершины не только наследуют права, но и получают их непосредственно.

4. Расширение и оптимизация RP-деревьев

В дальнейшем, выбранный подход к построению отображения RP будем указывать в названии RP-дерева.

Определение 7. Пусть T – RP-дерево. ***RP-характеристикой*** T называется спецификация, указывающая какой именно подход был применен при построении отображения RP:

1. Дерево T называется *таксономическим* (или *нетаксономическим*, или *охватным*), если при распределении прав был использован строго таксономический (или нетаксономический, или иерархический охватный) подход.

2. Если важно подчеркнуть, что было использовано листовое (классовое) распределение полномочий, то T – *листовое* (*классовое*) дерево.

Определение 8. ***Расширение*** RP-дерева T – это процесс построения RP-дерева T' такого, что

1. T является подграфом T'.
2. $\forall r \in R_T$: $RP_T(r) = RP_{T'}(r)$ – множество RP-классов T является подмножеством множества RP-классов T'.

Теорема 3. *Произвольное RP-дерево может быть расширено до таксономического (в общем случае классового) RP-дерева.*

Доказательство. Пусть T – произвольное RP-дерево. Построим искомое RP-дерево T'. Все вершины, дуги и полномочия дерева T перенесем в дерево T'. Тем самым условия определения 8 будут выполнены.

Пусть каждой листовой вершине r_i сопоставлен набор полномочий $P_i = \{p_{i1}, \ldots, p_{im_i}\}$. Если $|P_i| = m_i > 1$, то в дереве T' к этой вершине присоединим m_i листовых вершин, каждая из которых будет наделена правом p_{ij} $(j \in \{1, \ldots, m_i\})$.

Двигаясь по дереву T' от листьев к корню, каждую нелистовую вершину r пополним сыновьями-листьями по числу полномочий из набора $RP(r)$ дерева T, которые не были унаследованы (каждой новой вершине припишем соответствующее право).

В результате, в дереве T' каждая нелистовая вершина не получает ни одного полномочия непосредственно, а лишь наследует их от сыновей:

$$\forall r \notin R_L : RP(r) = \bigcup_{r' \in Ch(r)} RP(r'). \tag{19}$$

Каждой листовой вершине дерева T' приписано одно единственное полномочие. Объединяя листовые вершины с одним и тем же значением $RP(r) = \{p_i\}$ в один класс разбиения листовых вершин $R_L^{(i)}$, получаем:

$$\forall r \in R_L^{(i)} : RP(r) = \{p_i\} = P_i,\ \forall i, j \in \{1, \ldots, k\} : P_i \cap P_j = \varnothing. \tag{20}$$

Итак, отображение RP удовлетворяет всем требованиям строго таксономического классового подхода, следовательно, T' – таксономическое RP-дерево (см. рис. 1). ■

Расширяя RP-дерево, мы, тем самым, строим новую ролевую политику, наследующую все роли и их иерархию из исходной модели.

В силу теоремы 3, одним из преимуществ классового распределения полномочий является возможность расширения нетаксономических или охватных RP-деревьев до строго таксономических классовых, то есть возможна смена произвольной RP-характеристики дерева на таксономическую. Но, к сожалению, при таком преобразовании, как правило, увеличивается количество ролей (и RP-классов) в системе.

В противовес расширению RP-дерева можно рассматривать в некотором смысле обратную операцию. Если в RP-дереве найдется хотя бы один RP-класс,

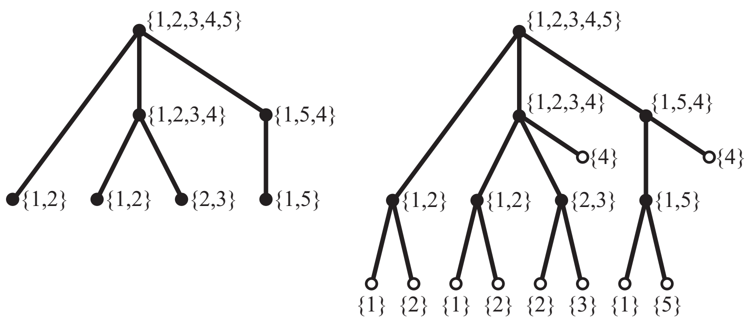


Рис. 1. Охватное классовое RP-дерево T (слева) расширено до таксономического классового RP-дерева T' (справа)

содержащий несколько ролей, то дерево не оптимально, а это свидетельствует о наличии в политике безопасности «дублирующих» ролей. Естественно попытаться преобразовать иерархию ролей так, чтобы результирующее RP-дерево стало оптимальным и при этом не изменилось множество RP-классов системы.

Определение 9. Два RP-дерева T и T' ***эквивалентны***, если множества их RP-классов совпадают (совпадают различные наборы полномочий, встречающиеся в структуре).

Например, если в RP-дереве T на рисунке 1 удалить левое ребро, то получим RP-дерево, задающее уже другую иерархию ролей, но эквивалентное исходному.

Определение 10. ***Оптимизация*** RP-дерева T – это процесс построения RP-дерева T' такого, что

1. T' эквивалентно T.

2. T' – оптимальное RP-дерево.

Заметим, что RP-дерево, полученное в результате оптимизации, будет листовым в силу оптимальности.

Попытаемся ответить на следующие вопросы. Любое ли RP-дерево поддается оптимизации? Как при этом ведет себя RP-характеристика дерева?

Если RP-дерево является листовым и вершины в пределах одного RP-класса не связаны дугами (иначе достаточно произвести попарное *стягивание* таких вершин, как эта операция понимается в теории графов [3]), то добиться оптимальности в ряде случаев можно за счет перестройки древовидной структуры и изменения RP-характеристики на охватную. Этот подход не столь интересен, так как, исходя из практических приложений, желательно получить эквивалентное оптимальное таксономическое RP-дерево.

5. Ролевая политика безопасности на произвольном ориентированном графе

Получение эквивалентной оптимальной структуры с той же RP-характеристикой представляется возможным за счет отказа от древовидности и построения эквивалентного ориентированного графа, задающего иерархию ролей.

Теорема 4. *Ориентированный граф задает иерархию ролей (является орграфом ролей) тогда и только тогда, когда в нем отсутствуют ориентированные циклы.*

Доказательство. Отсутствие ориентированных циклов необходимо и достаточно для существования отношения частичного порядка, а именно свойств транзитивности и антисимметричности. ■

Заметим, что в орграфе без ориентированных циклов найдется как минимум один сток (вершина с нулевой полустепенью исхода: $d^-(t_i) = 0$), и как минимум один источник (вершина с нулевой полустепенью захода: $d^+(s) = 0$). Далее будем рассматривать ориентированные графы с одним источником.

Распределение прав по произвольнму орграфу ролей, также как и по дереву ролей, может производится одним из трех способов. При этом построение отображения RP начинается либо со стоков (листовое распределение) либо с классов эквивалентности, на которые разбиты стоки (классовое распределение).

Определения оптимальности, расширяемости, эквивалентности и оптимизации очевидным образом переносятся на случай RP-орграфа (помеченного орграфа ролей).

Теорема 5. *Произвольное RP-дерево может быть оптимизировано до RP-орграфа.*

Доказательство. В RP-дереве достаточно *склеить* вершины, соответствующие эквивалентным ролям, если они не соединены дугами, либо попарно *стянуть*, если такие дуги имеются (операции склейки и стягивания вершин понимаются в соответствии с определениями теории графов [3]). В результате, множество RP-классов останется прежним, но орграф будет оптимальным (см. рис. 2). ■

Следствие 5.1. *Из алгоритма построения эквивалентного оптимального RP-орграфа G непосредственно следует ряд свойств этой структуры:*

1. *G имеет один источник s.*

2. *Число стоков t_i в G совпадает с числом классов разбиения $R_L^{(i)}$ листовых вершин исходного RP-дерева T.*

3. *Если исходное RP-дерево T являлось оптимальным, то $G = T$.*

4. *G – листовой RP-орграф.*

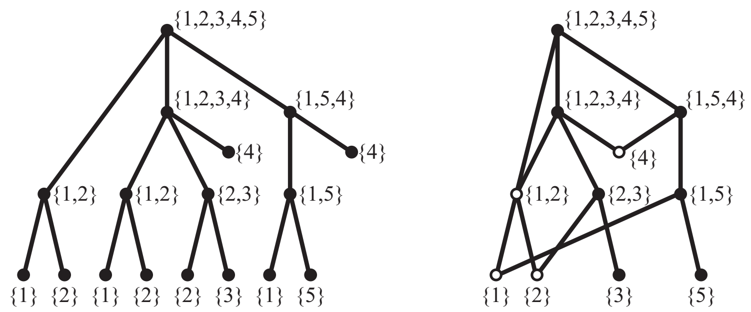


Рис. 2. Таксономическое классовое RP-дерево T (слева) оптимизировано до листового таксономического RP-орграфа G (справа)

Следствие 5.2. *Если исходное RP-дерево таксономическое, то построенный по предложенному алгоритму эквивалентный оптимальный RP-орграф также таксономический.*

Доказательство. Стягивание двух вершин, соответствующих эквивалентным ролям, по дуге их соединяющей не изменяет RP-характеристику. При склейке вершин из одного RP-класса, результирующий набор сыновей будет распределен по тем же RP-классам, что и в исходном RP-дереве – тем самым сохранится таксономичность структуры. ■

Следствие 5.3. *Если исходное RP-дерево нетаксономическое, то построенный по предложенному алгоритму эквивалентный оптимальный RP-орграф также нетаксономический.*

Следствие 5.4. *Если исходное RP-дерево охватное, то построенный по предложенному алгоритму эквивалентный оптимальный RP-орграф может оказаться охватным, нетаксономическим или таксономическим.*

Обобщая вышесказанное, получаем следующую возможную последовательность построения ролевой политики безопасности:

1. Исходя из содержательной постановки задачи, построить RP-дерево T_1 (листовое или классовое).

2. Расширить T_1 до таксономического (в общем случае классового) RP-дерева T_2 (см. теорему 3).

3. Преобразовать T_2 в эквивалентный оптимальный таксономический RP-орграф T_3 (см. теорему 5).

Таким образом, любую ролевую модель распределения полномочий можно расширить до политики, в которой иерархия ролей задана орграфом без ориентированных циклов, роли распределены в соответствии со строго таксономическим листовым подходом и RP-структура оптимальна.

Оказывается, предложенный в теореме 5 алгоритм обратим: по произвольному RP-орграфу можно построить эквивалентное (но не обязательно оптимальное) RP-дерево.

Теорема 6. *Для произвольного RP-орграфа существует эквивалентное ему RP-дерево.*

Доказательство. Пусть дан RP-орграф G. Будем строить эквивалентную ему RP-структуру T. На первом шаге каждому стоку t_i орграфа G сопоставляем $d^+(t_i)$ листьев в T («оригинал» и $(d^+(t_i)-1)$ «дублей»). Эта операция называется *расщеплением* вершины (если полустепень захода равна единице, то имеется только «оригинал»).

Далее, двигаясь по орграфу G от нижних ярусов к источнику, последовательно расщепляем все вершины. «Оригинал» и «дубли» наделяем теми же правами, что были у вершины их образующей. К «оригиналу» присоединяем уже существующие вершины структуры T из тех, что не имеют входящих дуг, восстанавливая сыновей расщепляемой вершины орграфа G (такие вершины в T всегда найдутся по построению). К каждому «дублю» добавляем вершины и дуги так, чтобы подграф, порожденный «дублем», представлял собой копию поддерева, порожденного «оригиналом».

Очевидно, что построенная таким образом иерархия T является RP-деревом и задает те же RP-классы, что и исходный RP-орграф G, то есть ему эквивалентна (см. рис. 3). ■

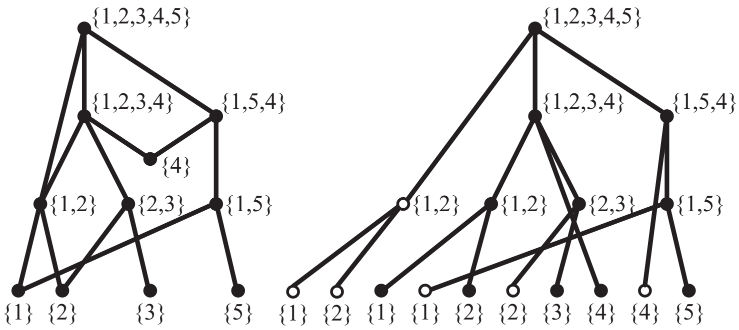


Рис. 3. RP-орграф G (слева) и эквивалентное ему RP-дерево T (справа)

Следствие 6.1. *Количество вершин RP-дерева T, эквивалентного RP-орграфу G и построенного по алгоритму, описанному в теореме, равно*

$$\sum_{r\in R_G} \left(1 + (d^+(r) - 1)|R_{T(r)}|\right), \qquad (21)$$

где R_G – множество вершин орграфа G, $R_{T(r)}$ – множество вершин поддерева, порожденного той вершиной дерева T, которая соответствует вершине r орграфа G.

Заметим, что теорема 6 дает возможность свести исследование ролевой политики безопасности на произвольном RP-орграфе к изучению эквивалентного RP-дерева.

Таким, образом, теоремы 5 и 6 позволяют выполнять различные эквивалентные преобразования иерархии ролей в зависмости от того, какой признак более значим: древовидность или оптимальность.

Литература

1. Гайдамакин Н.А. Разграничение доступа к информации в компьютерных системах / Н.А. Гайдамакин. – Екатеринбург: Изд-во Урал. ун-та, 2003. – 328 с.
2. Девянин П.Н. Модели безопасности компьютерных систем / П.Н. Девянин. – М.: Издательский центр «Академия», 2005. – 144 с.
3. Новиков Ф.А. Дискретная математика для программистов / Ф.А. Новиков – СПб.: Питер, 2001. – 304 с.

Математические структуры и моделирование
2009, вып. 19, с. 97–103

УДК 65.012.810

ИССЛЕДОВАНИЕ БЕЗОПАСНОСТИ КОМПЬЮТЕРНЫХ СИСТЕМ В МОДЕЛИ ДИСКРЕЦИОННОГО РАЗДЕЛЕНИЯ ДОСТУПА HRU

Д.М. Бречка, С.В. Белим

Исследуется возможность расширения классической модели дискреционной безопасности HRU, разделение доступа в которой основывается на матрице доступов. Основной целью является выявление систем, удовлетворяющих требованиям безопасности.

Введение

Фундаментальным понятием в сфере защиты информации компьютерных систем является политика безопасности. Под ней понимают совокупность норм и правил, регламентирующих процесс обработки информации, выполнение которых обеспечивает состояние защищенности информации в заданном пространстве угроз. Формальное выражение политики безопасности называют моделью безопасности.

Среди программно-технических методов защиты информации выделяют в первую очередь разграничение доступа. Разграничение доступа непосредственно обеспечивает конфиденциальность информации, а также снижает вероятность реализации угроз целостности и правомерной доступности.

Под разграничением доступа к информации в компьютерных системах понимают разделение информации, циркулирующей в системе, на части, элементы, компоненты, объекты и т. д., и организацию такой системы работы с информацией, при которой пользователи имеют доступ только и только к той части (к тем компонентам) информации, которая им необходима для выполнения своих функциональных обязанностей или необходима исходя из иных соображений [3].

Наиболее распространенной и исторически первой является дискреционная политика безопасности. Дискреционная политика описывает отношения между субъектами и объектами компьютерной системы на основе прав, которые субъекты имеют над объектами. При этом под субъектами понимаются активные сущности системы, которые могут изменять состояние системы через порождение процессов над объектами, в том числе, порождать новые объекты. Объекты

Омский государственный университет им. Ф. М. Достоевского.
E-mail: dbrechkawork@yandex.ru

— пассивные сущности системы. В дискреционных моделях исследуеются состояния системы на выявление возможности угроз информации [2].

1. Модель HRU

Модель Харисона-Руззо-Ульмана (HRU) является классическим примером дискреционной модели безопасности. Согласно этой модели компьютерная система представляется набором следующих сущностей:

— множество объектов (O) — пассивные сущности системы;

— множество субъектов (S) — активные сущности;

— права доступа (R) — действия, которые субъект может производить над объектом.

Определение 1. Матрица доступа — это таблица, в строках которой расположены субъекты системы, а в столбцах — объекты. Каждая ячейка матрицы доступа специфицирует права доступа субъектов к объектам.

Возможности изменения матрицы доступа определяется шестью примитивными операторами:

$-Enter\ r\ into\ A[s,o]$ — внести право r в ячейку $A[s,o]$;

$-Delete\ r\ from\ A[s,o]$ — удалить право r из ячейки $A[s,o]$;

$-Create\ subject\ s$ — создать субъект s (т. е. новую строку матрицы A);

$-Create\ object\ o$ — создать объект o (т. е. новый столбец матрицы A);

$-Destroy\ subject\ s$ — уничтожить субъект s;

$-Destroy\ object\ o$ — уничтожить объект o.

В результате выполнения примитивного оператора осуществляется переход системы из состояния $Q = (S, O, A)$ в новое состояние $Q' = (S', O', A')$.

Из примитивных операторов сторятся команды, состоящие из двух частей:

— условие, при котором выполняется команда;

— последовательность примитивных операторов.

Таким образом команда имеет вид:

$$
\begin{array}{l}
Command\ c(x_1,\ x_2,\ ...,\ x_k) \\
\quad if\ r_1\ in\ A[x_{s_1},\ x_{o_1}]\ and \\
\quad\ \ r_2\ in\ A[x_{s_2},\ x_{o_2}]\ and \\
\quad\ \ . \\
\quad\ \ . \\
\quad\ \ . \\
\quad\ \ r_m\ in\ A[x_{s_m},\ x_{o_m}] \\
\quad then\ op_1,\ op_2,\ ...,\ op_n
\end{array}
$$

Каждое состояние системы Q_i является результатом выполнения некоторой команды c, применимой, согласно ее условиям, к предыдущему состоянию Q_{i-1}, и определяет отношения доступа, которые существуют между сущностями системы.

Безопасность системы определяется некоторыми условиями на начальное состояние системы Q_0, а также особенностями системы команд

Определение 2. Система является безопасной относительно права r, если для заданного начального состояния $Q_0 = (S_0, O_0, A_0)$ не существует применимой к Q_0 последовательности команд, в результате которой право r будет занесено в ячейку матрицы $A[s, o]$, в которой оно отсутствовало в начальном состоянии Q_0.

Однако задача проверки данного критерия на истинность для произвольной системы алгоритмически неразрешима [1].

2. Примитивный базис модели HRU

Обозначим множество всех возможных матриц доступа через M. Команды HRU будут являться отображениями на этом множестве. Выделим базис системы команд HRU, для этого введем понятие базисного набора операторов.

Определение 3. Базисный набор операторов — это такой набор операторов, что любую команду HRU можно представить в качестве последовательности операторов из данного набора.

Теорема 1. *Примитивные операторы модели HRU образуют базисный набор.*

Доказательство. Покажем полноту системы примитивных операторов.

Так как работа команд HRU направлена исключительно на изменение матрицы доступа, покажем, что при помощи системы примитивных операторов можно построить любую матрицу доступа, то есть построить любую команду.

Рассмотим матрицы A и A' принадлежащие множеству всех матриц доступа M. Преобразуем A в A' при помощи системы примитивных операторов. Для этого:

1. Применим оператор *Destroy object* к матрице A до тех пор, пока количество объектов в матрице A не станет равным нулю.

$For\ i = m\ to\ 0$

$Destroy\ object\ o_i$,

где m — количество объектов в матрице A.

2. Применим оператор *Destroy subject* к матрице A до тех пор, пока количество субъектов в матрице A не станет равным нулю.

$For\ i = n\ to\ 0$

$Destroy\ subject\ s_i$,

где n — количество субъектов в матрице A.

3. Применим оператор *Create subject* к матрице A до тех пор, пока количество субъектов в матрице A не станет равным количеству субъектов в матрице A'.

$For\ i = 0\ to\ n'$

$Create\ subject\ s_i$,

где n' — количество субъектов в матрице A'.

4. Применим оператор *Create object* к матрице A до тех пор, пока количество объектов в матрице A не станет равным количеству объектов в матрице A'.

$For\ i = 0\ to\ m'$

$Create\ object\ o_i$,

где m' — количество объектов в матрице A'.

5. Для каждого права $r \subseteq R$ применим оператор $Enter\ r\ into\ A[s, o]$ если право r содержится в соответствующей ячейке $A'[s', o']$.

$For\ i = 0\ to\ R$

$For\ j = 0\ to\ m$

$For\ k = 0\ to\ n$

$If\ r_i\ in\ A'[j, k]$

$Enter\ r_i\ into\ A[j, k]$,

где R — количество всех возможных прав, m — количество объектов, n — количество субъектов.

Таким образом, при помощи примитивных операторов, произвольно выбранная матрица $A \subseteq M$, может быть перобразована в произвольно выбранную матрицу $A' \subseteq M$, что доказывает полноту системы. ■

Как видно из доказательства теоремы, для построения полной системы используется только пять из шести примитивных операторов.

Следствие 1.1. *Оператор delete r можно представить в виде последовательности операторов enter r, create object o, create subject s, destroy object o, destroy subject s.*

Доказательство. Рассмотрим матрицы B и B' принадлежащие множеству всех матриц доступа M такие, что количество строк и количество столбцов в этих матрицах равны. В какой-либо ячейке матрицы B' отсутствует право r_j, которое присутствует в соответствующей ячейке матрицы B. Все остальные ячейки матриц B и B' полностью идентичны. Преобразуем матрицу B в B' способом, аналогичным описанному в доказательстве теоремы 1. Данный способ позволяет преобразовать друг в друга произвольно выбранные матрицы, но не использует оператор *delete r*. Следовательно, для удаления права из матрицы B можно обойтись без оператора *delete r*. Это доказывает, что оператор *delete r* является избыточным и представляется в виде последовательности операторов *enter r*, *create object o*, *create subject s*, *destroy object o*, *destroy subject s*. ■

Введем понятие минимального набора операторов.

Определение 4. Минимальный базисный набор — это такой набор операторов, что ни один оператор данного набора не может быть представлен в качестве последовательности операторов из данного набора.

Теорема 2. *Набор примитивных операторов enter r, create object o, create subject s, destroy object o, destroy subject s является минимальным.*

Доказательство. Согласно определению, набор операторов не будет минимальнм, если какой-либо оператор из набора можно представить в виде набора других операторов принадлежащих даннму набору. Допустим, что рассматриваемый набор не минимален. Тогда для преобразования произвольно выбранной матрицы $A \subseteq M$ в произвольно выбранную матрицу $A' \subseteq M$ можно будет обойтись без какого-либо оператора. Однако, как показано в доказателстве теоремы 1, для преобразования матриц необходимы все операторы рассамтриваемого набора. Следовательно рассматриваемый набор будет являться минимальным. ■

3. Построение базиса, отличного от примитивного

Рассмотрим возможность перехода к новому базису: представим какой-либо из операторов, входящих в базисный набор в виде последовательности операторов из другого набора.

Пусть S - множество субъектов системы, O - множество объектов, R - множество прав доступа. A - матрица доступа. Представим каждую ячейку матрицы A в виде последовательности бит таким образом, что каждому биту будет соответствовать какое-либо право из доступного набора прав R. Количество бит будет зависить от количества возможных прав, то есть если $|R| = l$, тогда для кодирования всех прав понадобится l бит. Наличие или отсутствие какого-либо бита означает наличие или отсутствие соответствующего права в данной ячейке.

Представим операцию добавления права в ячейку виде последовательности логических функций. Будем применять логические функции к матрице доступа поэлементно, то есть для изменения права доступа субъекта s на объект o необходимо изменить ячейку $A[s, o]$.

Рассмотрим операцию добавления права $r_k \in R$ субъекту s на объект o. Для ячейки матрицы $A[s, o]$ необходимо будет выполнить логичесукю операцию «OR» с таким двоичным числом, в котором бит с номером k равен 1, а все остальные биты равны 0.

Логическая операция «OR» может быть выражена через логические операции «AND» и «NOT»: $OR(a, b) = NOT(AND(NOT(a), NOT(b)))$. Следовательно, операцию добавления права в ячейку $A[s, o]$ можно представить в виде двух логических операций «AND» и «NOT».

Таким образом можно получить новый базис, отличный от примитивного, он будет состоять из операторов *create object* o, *create subject* s, *destroy object* o, *destroy subject* s, $AND(a, b)$ и $NOT(a)$.

4. Монооперационные системы

Как уже говорилось ранее, для произвольной системы не существует алгоритма, проверяющего является ли ее исходное состояние безопасным. Для рассмотрения условий, при которых такой алгоритм существует вводится понятие монооперационной системы.

Определение 5. Система называется монооперационной, если каждая команда данной системы выполняет один примитивный оператор.

Теорема 3. *Существует алгоритм, проверяющий: является ли исходное состояние монооперационной системы безопасным по отношению к праву* $r \in R$.

Доказательство приводится в [1].

Расширим классическое понятие монооперационной системы, с учетом базисного набора операций.

Определение 6. Монооперационная система в базисе B — это такая система, каждая команда которой содержит ровно один оператор из базиса B.

Формулировку теоремы 3 также можно сделать более общей.

Теорема 4. *Существует алгоритм, проверяющий: является ли исходное состояние монооперационной системы в базисе* B *безопасным по отношению к праву* $r \in R$.

Доказательство. Необходимо показать, что число последовательностей команд монооперационной системы в базисе, которые необходимо проверить, ограничено и сами команды имеют конечную длину. В этом случае алгоритмом проверки безопасности будет являться алгоритм перебора всех последовательностей команд и проверки их конечного состояния на отсутствие утечки права r.

Нет необходимости рассматривать команды, содержащие операторы *delete...* и *destroy...*, так как необходимо проверить наличие права после выполнения команды, а не его отсутствие. Также нет необходимости рассматривать последовательности, содержащие более одного оператора *create...*, так как все последовательности, которые проверяют или вносят права в новые элементы матрицы, могут быть заменой параметров в командах, представленных последовательностями, действующими с существующими субъектами и объектами. Одна команда создания субъекта необходима, если в начальном состоянии системы субъекты отсутствуют. Таким образом, необходимо рассмотреть только последовательности команд, которые содержат операторы *enter r* и один оператор *create subject*. Число различных операторов *enter r* можно вычичлить следующим образом:

$n = |R|(|S_0| + 1)(|O_0| + 1)$.

Все команды, содержащие один и тот же оператор, но разные условия, объединяются в одну команду с составным условием.

Таким образом, число последовательностей команд, которые необходимо проверить равняется $n!$, при этом длина каждой команды равна n. ■

Непредсказуемость сложных систем, является одним из главных недостатков модели HRU. Наложение условия монооперационности значительно сужает класс безопасных систем. При введении понятия базиса монооперационной системы, появляется возможноть рассматривать более широкий класс систем, которые также будут являться безопасными. Для того, чтобы проверить является ли компьютерная система безопасной, достаточно найти такой базис, в котором она будет монооперационна.

Литература

1. Девянин П.Н. Модели безопасности компьютерных систем / П.Н. Девянин – М.: издательский центр «Академя», 2005.
2. Гайдамакин Н.А. Разграничение доступа к информации в компьютерных системах / Н.А. Гайдамакин – Уральский университет, 2003.
3. Грушо А.А. Тимонина Е.Е. Теоретические основы защиты информации /А.А. Грушо, Е.Е. Тимонина – Агентство «Яхтсмен», 1996.

Математические
структуры и моделирование
2009, вып. 19, с. 104–107

УДК 004.056.5:005.311.7

ТЕОРЕТИКО-ИГРОВОЙ ПОДХОД К ВЫБОРУ ОПТИМАЛЬНЫХ СТРАТЕГИЙ ЗАЩИТЫ ИНФОРМАЦИОННЫХ РЕСУРСОВ

Т.В. Вахний, А.К. Гуц

В данной работе для поиска наиболее оптимальных стратегий защиты информационных ресурсов используется математическая игра двух сторон, одной из которых является система защиты компьютерной информации, а с другой – атаки азартных хакеров. Применение игровых методов дает преимущества администратору безопасности перед субъективными случайными решениями и обеспечивает оптимизацию стратегий защиты компьютерной информации. Учет психологии азартного хакера позволяет направлять его активность в ложном направлении.

Введение

В настоящее время, учитывая широкое распространение информационных систем, интегрированных в глобальные информационно-вычислительные сети, приходится опасаться удаленных атак хакеров. Одной из существенных особенностей обеспечения защиты информационных ресурсов является недостаточность информации о возможных атаках, времени их проведения и их последствиях. Поэтому задачи обеспечения защиты информационных ресурсов следует относить к «задачам о выборе решений в условиях неопределенности».

В данной работе для поиска наиболее оптимальных стратегий защиты информационных ресурсов используется математическая игра двух сторон, одной из которых является система защиты компьютерной информации, а с другой – возможные атаки азартных хакеров. При составлении матрицы игры можно считать, что хакер увлечен желанием нанести как можно больший ущерб атакуемой компьютерной системе. Цель администратора безопасности в матричной игре состоит в том, чтобы позволить хакеру причинить наименьший ущерб.

Идея использования теоретико-игрового подхода в теории защиты информационных ресурсов не является новой [1-3], но соответствующие работы малодоступны. Поэтому имеет смысл уточнить постановку задачи и ход ее решения.

Омский государственный университет.
E-mail: `vahniytv@mail.ru`

В результате математического моделирования игры можно оценить эффективность стратегий администратора безопасности по защите информационных ресурсов и выбрать из них наиболее эффективные.

Постановка и решение задачи

Будем понимать стратегии хакера как строки x_i $(i = 1, 2, ..., n)$ некоторой матрицы, а стратегии администратора информационных ресурсов – как ее столбцы y_j $(j = 1, 2, ..., m)$. К стратегиям хакера можно отнести различные виды компьютерных атак. Например, это могут быть удаленное или локальное проникновение в компьютер, удаленное или локальное блокирование компьютера, применение сетевых сканеров для сбора информации о компьютерах сети и программах потенциально уязвимых к атакам, использование сканеров уязвимых мест программ в поисках компьютеров, уязвимых к тому или иному конкретному виду атаки, применение вскрывателей паролей, применение сетевых анализаторов (снифферов) и др.

К стратегиям администратора можно отнести различные варианты использования методов и средств защиты информации. Например, применение и регулярное обновление антивирусных программ, шифрование, использование межсетевых экранов и средств обнаружения атак, оперативная установка от производителей исправлений для программ (чтобы ликвидировать неблагоприятные последствия ошибок в них), применение вскрывателей паролей и сканеров уязвимых мест и др.

Для проведения на компьютере игры A надо также знать результаты игры a_{ij} при каждой паре стратегий x_i и y_j (например, a_{ij} – причиненный материальный ущерб) и вероятности реализации атак хакеров $p(x_i)$ при выбранной стратегии x_i. Построив игровую матрицу (см. табл. 1) и проанализировав ее, можно заранее оценить затраты каждого решения по защите компьютерной информации и рекомендовать наиболее эффективные варианты для всего диапазона атак.

Таблица 1.

		y_1	y_2	...	y_m
x_1	$p(x_1)$	a_{11}	a_{12}	...	a_{1m}
x_2	$p(x_2)$	a_{21}	a_{22}	...	a_{2m}
...	...	...	...	...	...
x_n	$p(x_n)$	a_{n1}	a_{n2}	...	a_{nm}

Если построена игровая матрица (a_{ij}), в которой результатами игры являются материальные потери от атак, то наилучшей в условиях имеющейся информации об атаках будет стратегия системы защиты компьютерной информации y_j, при которой будут минимальны средние потери, т. е. будет минимальна сумма [1–3]:

$$\sum_{i=1}^{n} a_{ij} \cdot p(x_i).$$

Вероятности реализации атак $p(x_i)$ могут быть определены по результатам статистических исследований. Если вероятности атак неизвестны, то предполагается, что все они равновероятны, т. е. $p(x_i) = 1/n$.

Азартный хакер увлечён желанием нанести как можно больший ущерб атакуемой компьютерной системе. В силу своей психологии он преувеличивает свои выигрыши и преуменьшает свои неудачи в предыдущих попытках атак на систему, воспринимая игру A как матричную игру $f(A)$ с матрицей $(f(a_{ij}))$, где f — так называемая функция полезности. В случае азартного нарушителя эта функция задается непрерывной выпуклой (вниз) вещественной функцией $f : \mathbb{R} \to \mathbb{R}$ [3, с. 222]. На рис. 1 приводится вид функции полезности азартного хакера. Такой функцией может являться, например, функция $f(a) = e^a - 1$.

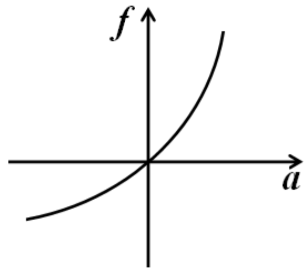


Рис. 1. Вид функции полезности азартного игрока [3, с. 222]

Обозначим через $val(A)$ значение матричной игры A с матрицей (a_{ij}). В случае азартной функции полезности имеют место утверждения [3]:

1) из $val(A) = 0$ следует $val(f(A)) \geq 0$, т. е. нарушитель может видеть победу там, где её нет;

2) из $val(A) > 0$ следует $val(f(A)) \geq val(A)$, т. е. азартный нарушитель преувеличивает размер успеха;

3) при любом опыте l предыдущих вторжений существует такая игра A_0, что $val(A_0) < 0$ (реальный проигрыш, неудачная атака) и $val(A_0 + lE) > f(l)$, где E — матрица, состоящая из единиц, т. е. азартный нарушитель всегда будет повторять некоторые проигрышные атаки (игру A_0).

Учет психологии азартного хакера и моделирование его поведения позволяет строить ловушки либо для его идентификации, либо для направления его активности в ложном направлении.

Оценка материальных потерь

Построение игровых матриц и выбор наиболее приемлемых решений при использовании игровых моделей требует оценки результатов функционирования систем защиты компьютерной информации в целом при различных возможных вариантах решений. Опишем один из способов определения коэффициентов a_{ij} матрицы игры A.

Единичные потери P_{ij}^1 при взломе j-ой рабочей станции в случае однократной реализации угрозы x_i, можно оценить следующим образом:

$$P_{ij}^1 = R_j k_i,$$

где R_j – стоимость ресурса «рабочая станция пользователя» при использования j-ой комбинации методов и средств защиты; k_i – процент потерь в случае реализации угрозы x_i на данном ресурсе. Стоимость ресурса R_j обычно включает стоимость сопровождения и восстановления, прямые затраты на покупку и обновление соответствующего оборудования и программного обеспечения, расходы на поддержание информационной системы, административные расходы, затраты на обучение пользователей и убытки от вынужденных простоев.

Годовая оценка инцидента N_i, т. е. число, отражающее частоту проявления угрозы x_i в год, может быть рассчитана так:

$$N_i = s\nu_i,$$

где s – число подверженных атаке рабочих станций и ν_i – частота реализации угрозы x_i в год (может быть найдена на основе собственного опыта или усредненной статистической информации).

Годовые потери P_{ij} j-ой рабочей станции в результате реализации угрозы x_i можно оценить следующим образом:

$$P_{ij} = P_{ij}^1 N_i.$$

В качестве коэффициентов a_{ij} матрицы игры A можно рассматривать годовые потери P_{ij} для всех вариантов комбинаций x_i $(i = 1, 2, ..., n)$ и y_j $(j = 1, 2, ..., m)$.

Заключение

Таким образом, применение игровых методов дает преимущества администратору безопасности перед субъективными случайными решениями и обеспечивает оптимизацию стратегий защиты компьютерной информации. Организация проигрышных атак и подробное исследование матричной игры A_0 сводится к изучению психологии азартного нарушителя.

Литература

1. Воробьев А.А. Методы оценивания и обеспечения гарантированного уровня защиты информации от несанкционированного доступа в вычислительной сети автоматизированной системы управления: Автореф. дис. … к-та техн. наук / А.А. Воробьёв. — СПб., 1997. – 15 с.
2. Нестеров С.А. Разработка методов и средств проектирования инфраструктуры обеспечения информационной безопасности автоматизированных систем: Автореф. дис. … к-та техн. наук / С.А. Нестеров. — СПб., 2002. — 18 с.
3. Матричные игры / Под. ред. Н.Н. Воробьева. – М: Государственное издательство физико-математической литературы, 1961. – 280 с.

Математические структуры и моделирование
2009, вып. 19, с. 108–113

УДК 65.012.810

МОДЕЛИРОВАНИЕ РОЛЕВОЙ ПОЛИТИКИ БЕЗОПАСНОСТИ В СООТВЕТСТВИИ СО СТАНДАРТОМ СТО БР ИББС-1.0-2008

Ю.С. Ракицкий, С.В. Белим

В статье рассматривается вопрос моделирования ролевой политики безопасности в соответствии с ее описанием в стандарте Банка России для организаций банковской системы Российской Федерации.

Введение

Анализ различных организационно - управленческих и организационно - технологических схем показывает, что, в реальной жизни сотрудники предприятий, учреждений выполняют определенные функциональные обязанности не от своего личного имени, а в рамках некоторой должности. Должность, которую можно трактовать как определенную роль, представляет некоторую абстрактную, точнее обобщенную сущность, выражающую определенный тип функций и тип положения работника (подчиненность, права и полномочия). Таким образом, в реальной жизни в большинстве организационно-технологических схем права и полномочия предоставляются конкретному сотруднику не лично (непосредственно), а через назначение его на определенную должность (роль), с которой он и получает некоторый типовой набор прав и полномочий. Ролевое разграничение доступа является развитием политики дискреционного разграничения доступа, при этом права доступа субъектов системы (т.е. сотрудников предприятия, занимающих определенную должность) на объекты с учетом специфики их применения, образуя роли.

Ярким примером описанных предприятий являются коммерческие банки. Учитывая законодательство в области банковской деятельности (ст. 26 «Банковская тайна» закона «О банках и банковской деятельности» и закон «О персональных данных»), проблема обеспечения информационной безопасности компьютерных систем в организациях банковской системы является весьма актуальной. Центральный Банк Российской Федерации (Банк России) выпустил серию документов, посвященных обеспечению информационной безопасности

Омский государственный университет им. Ф.М. Достоевского.
E-mail: `yrakitsky@rambler.ru`

организаций банковской системы. Одним из таких документов является стандарт Банка России СТО БР ИББС-1.0-2008 «Обеспечение информационной безопасности организаций банковской системы Российской Федерации». Авторы данного стандарта предлагают при построении политики информационной безопасности определить и разграничить роли сотрудников банка.

1. Описание политики информационной безопасности в стандарте СТО БР ИББС-1.0-2008

Согласно пункту стандарта 7.2.1 «Роль – это заранее определенная совокупность правил, устанавливающих допустимое взаимодействие между субъектом, например, сотрудником организации, и объектом, например, программно - аппаратным средством. Для эффективного выполнения целей организации и задач по управлению активами должны быть выделены и определены соответствующие роли персонала организации». Таким образом, в любой автоматизированной компьютерной системе предоставление доступа должно осуществляться в соответствии с ролевой моделью разграничения доступа.

Согласно пункту стандарта 7.2.3 «Не рекомендуется, чтобы одна персональная роль целиком отражала цель, например, включала все правила, требуемые для реализации бизнес-процесса. Совокупность правил, составляющих роли, не должна быть критичной для организации с точки зрения последствий успешного нападения на ее исполнителя. Не следует совмещать в одном лице (в любой комбинации) роли разработки, сопровождения, исполнения, администрирования или контроля, например, исполнителя и администратора, администратора и контролера или других комбинаций». Таким образом, выделяются роли исполнителей, контролеров, администраторов и сопровождения, которые должны в какой-либо комбинации присутствовать в любом процессе в каждой автоматизированной компьютерной системе.

Согласно пункту стандарта 7.2.4 «Роль должна быть обеспечена ресурсами, необходимыми и достаточными для ее исполнения». Следовательно, любая роль не должна содержать избыточных прав доступа в автоматизированной компьютерной системе, то есть не должна обладать доступом к объектам, которые не используются при исполнении данной роли. Например, сотрудник банка, оформляющий в автоматизированной системе заявки на выдачу кредита клиенту банка не должен иметь доступ к информации о принятии вкладов от населения, но при этом должен обладать правами доступа на просмотр и редактирование анкетных данных клиентов, подавших заявку на получение кредита.

Приведенные выше требования являются частью общих требований по обеспечению информационной безопасности при назначении и распределении ролей и обеспечении доверия к персоналу. На основании этих требований можно сформулировать формальную модель ролевой политики безопасности, в соответствии со стандартом СТО БР ИББС-1.0-2008.

2. Формализация политики безопасности

Базовая модель ролевого разграничения доступа включает в себя следующие множества: U — множество пользователей, R -– множество ролей, P – множество прав на работу в системе. Важную роль играет отображение

$$PA: \ R \longrightarrow 2^P, \tag{1}$$

определяющее множество прав доступа для заданной роли, при этом для каждого $p \in P \ \exists \ r \in R$ такая, что $p \in PA(r)$.

Для введения в модель контролирующих функций необходимо множество ролей R, которые в дальнейшем будем называть исполнительскими, дополнить множеством административных ролей ACR и множеством контролирующих ролей CR. При этом

$$R \cap ACR = \oslash, ACR \cap CR = \oslash, R \cap CR = \oslash \tag{2}$$

Также введем дополнительные множества: ACP — множество прав для административных ролей, CP — множество прав для контролирующих ролей. Множества P, CP и ACP также не имеют общих элементов. ACR осуществляют администрирование контролирующих ролей.

Для каждого права $p \in P$ должно быть определено множество контролирующих прав, обладание которыми необходимо для контроля над p. Введем соответствующее отображение

$$ControlRight: \ P \longrightarrow 2^{CP} \tag{3}$$

при этом $\forall p \in P \ ConrolRight(p) \neq \oslash$.

Также для любой роли должен существовать набор, контролирующих ролей, осуществляющих контроль над ней. Введем отображение

$$ControlRole: R \longrightarrow 2^{CR} \tag{4}$$

при этом

$$PA(r) = p_{i1}, p_{i2}, ..., p_{in} \Rightarrow PA(ControlRole(r)) = \bigcup ControlRight(p_{ij}). \tag{5}$$

Определение 1. В системе выполняются функции контроля, если в любой момент времени для любого $p \in P \ \exists \ cp_{i1}$, cp_{i2}, ..., $cp_{in} \subseteq CP$ такое, что cp_{i1}, cp_{i2}, ..., $cp_{in} \subseteq ControlRight(p)$, а для любой $r \in R \ \exists \ cr_{k1}$, cr_{k2}, ..., $cr_{kl} \subseteq CR$ такое, что cr_{k1}, cr_{k2}, ..., $cr_{kl} \subseteq ControlRole(r)$.

Аналогично, введем отображение

$$AdminRight: CP \longrightarrow 2^{ACP}, \tag{6}$$

при этом $\forall p \in P \ AdminRight(p) \neq \oslash$.

Введем отображение

$$AdminRole: \ CR \longrightarrow 2^{ACR}, \tag{7}$$

при этом

$$PA(cr) = cp_{i1}, cp_{i2}, ..., cp_{in} \Rightarrow PA(AdminRole(ar)) = ap_1, ap_2, ..., ap_m. \tag{8}$$

Определение 2. В системе выполняются функции администрирования, если в любой момент времени для любого $cp \in CP$ $\exists$ ap_{i1}, ap_{i2}, ..., $ap_{in} \subseteq ACP$ такое, что ap_{i1}, ap_{i2}, ..., $ap_{in} \subseteq AdminRight(cp)$, а для любой $cr \in CR$ $\exists$ ar_{k1}, ar_{k2}, ..., $ar_{kl} \subseteq ACR$ такое, что ar_{k1}, ar_{k2}, ..., $ar_{kl} \subseteq AdminRole(cr)$.

Аналогичным образом можно выделить в системе роли разработки и сопровождения.

3. Соответствие модели стандарту СТО БР ИББС-1.0-2008

Теперь покажем, что построенная модель соответствует требованиям стандарта СТО БР ИББС-1.0-2008.

Теорема 1. *Построенная модель удовлетворяет требованиям стандарта СТО БР ИББС-1.0-2008.*

Доказательство. Для доказательства теоремы необходимо показать, что введенные отображения соответствуют требованиям стандарта, и, наоборот, для каждого требования стандарта существует соответствующее отображение. Как было показано в пункте 2, в стандарте ролевой политике безопасности посвящено три пункта.

Согласно пункту 7.2.1 разграничение доступа должно производиться по ролевому принципу, что очевидно выполняется.

Согласно пункту 7.2.3 должны существовать исполнительские, административные, контролирующие роли, а также роли сопровождения, которые не должны совмещаться в одном лице. Это задается соотношением (2).

Согласно пункту 7.2.4 любая роль должна обладать необходимыми и достаточными правами на свое исполнение, что задается соотношениями (3), (4), (6), (7), согласно которым для любой роли существует контролирующая и административная роль, обладающая достаточными правами для исполнения своих функций. В то же время административная и контролирующая роли не обладают правами на исполнение других функций, что обозначено соотношением (2). ■

4. Пример построения модели

Для наглядности построим модель политики безопасности на основании введенных определений. Пусть задана иерархия исполнительских ролей, а также иерархия административных и контролирующих ролей.

Управляющий DIR является максимальной ролью в иерархии, минимальной ролью является служащий E. В каждом направлении деятельности определяется максимальная роль исполнительного директора TM, минимальной ролью направления является операционист O (рис. 1).

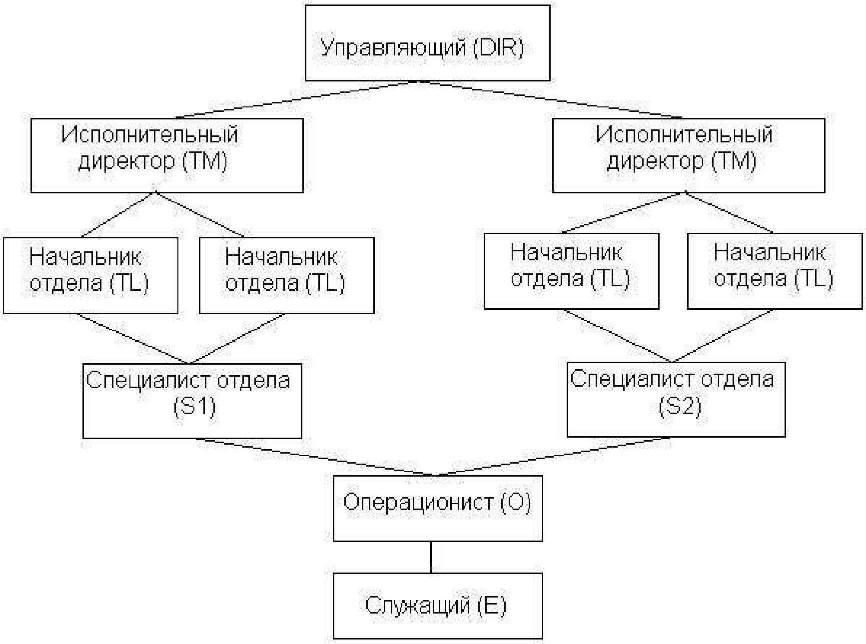


Рис. 1. Иерархия исполнительских ролей

Каждый из контролеров направления C_1, C_2 обеспечивают функции контроля за исполнителями начиная с начальников отделов, при этом функции контроля для C_1 и C_2 не пересекаются, т.е. C_1 не может контролировать направление деятельности 2 (рис. 2).

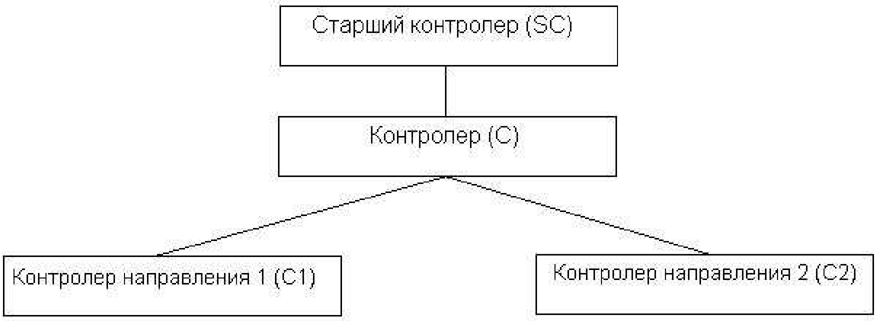


Рис. 2. Иерархия контролирующих ролей

Старший администратор A выполняет административные функции для SC и C, администраторы A_1 и A_2 выполняют административные функции по направлениям деятельности (рис. 3).

Ниже приведены таблицы, в которых описано соотношение между конторолирующими ролями и множествами сопоставленных им исполнительских ролей, а также между администраторскими правами и множествами сопоставленных им контролирующих ролей.

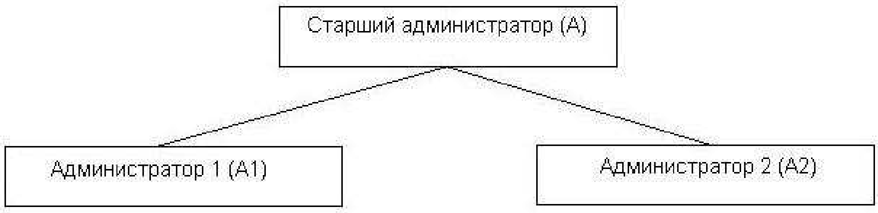


Рис. 3. Иерархия административных ролей

Контролирующая роль	Множество ролей
C_1	$[S_1, TM_1)$
C_2	$[S_2, TM_2)$
C	$[TM_1, TM_1]$
C	$[TM_2, TM_2]$

Административная роль	Множество ролей
A_1	$[C_1, C)$
A_2	$[C_2, C)$
A	$[C, C]$

Литература

1. Грушо А.А., Тимонина Е.Е. Теоретические основы защиты информации / А.А. Грушо, Е.Е. Тимонина. – М: Издательство Агенства Яхтсмен, 1996.
2. Зегжда Д.П., Ивашко А.М. Основы безопасности информационных систем / Д.П. Зегжда, А.М. Ивашко. – М.: Горячая линия – Телеком, 2000.
3. Щербаков А.Ю. Введение в теорию и практику компьютерной безопасности / А.Ю. Щербаков. – М.: Издатель Молгачева С.В., 2001.

Математические структуры и моделирование
2009, вып. 19, с. 114–121

УДК 004.4

ПРОБЛЕМА ВЫБОРА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ПРЕПОДАВАНИЯ БАЗОВЫХ ДИСИПЛИН И ПОДДЕРЖКИ УЧЕБНОГО ПРОЦЕССА НА ФКН ОМГУ

Д.Н. Лавров

В статье рассматриваются проблемы выбора легального программного обеспечения для преподования осноных дисциплн и обеспечения работы подразделений факультета компьютных наук Омского государственного университета.

Введение

Проблема выбора программного обеспечения одна из сложнейших проблем, особенно актуальной становиться в рамках преподавания компьютерных дисциплин. Огромное разнообразие предлагаемых программных продуктов различными производителями ставит разработчика учебного курса в условия множественного выбора.

С другой стороны необходимо обеспечить работу структурных подразделений, лабораторий, деканата и кафедр.

Какие инструменты изучать в первую очередь? Какими критериями при этом руководсвоваться?

Рассмотрим проблему выбора программного обеспечения (ПО) на примере факультета компьютерных наук (ФКН) Омского государственного университета. Выбор устанавливаемого ПО производиться ежегодно при подготовке компьютерных классов к новому учебному году. Причин радикально повлиявших на актуализацию этой проблемы несколько.

Во-первых, особое внимание к свободному и открытому программному обеспечению, проявляемое со стороны государственных органов. Стремление использовать его в образовательном процессе средней школы, попытки создания национальной операционной системы, – все это делает необходимым изучение этого программного обеспечения на факультете. Студенты должны влаеть не только проприетарными продуктами, но и лучшим свобобным и открытым ПО.

Омский государственный университет.
E-mail: `lavrov@omsu.ru`

Нужно отметить, что открытое ПО не является противопоставлением проприетарному ПО, как это часто явно или неявно преподносится в средствах массовой информации. Проприетарное ПО является лишь противопоставлением свободному ПО в определении Р.Столлмена [1]. Беспалатное ПО может быть проприетарным, открытое ПО также может быть проприетарным. Противопоставлением бесплатному ПО является коммерческое. Операционные системы Microsoft являются закрытыми, коммерческими продуктами, а Internet Explorer закрытым бесплатным продуктом. И то, и то являются проприетарным ПО. Операционная система Solaris является проприетарным продуктом, исходый код которой практически полностью открыт в проекте Open Solaris

Вторая причина — это усиление борьбы правоохранительных органов с использованием нелицензионного программного обеспечения. Это стимулирует к приведению использования ПО в вузе в соотвесвтие с законодательством. В частности это привело к оформлению MSDNAA-подписке на программное обеспечение фирмы Microsoft на ФКН. Эта частично решило проблему, но только частично, так как использование продуктов Microsoft по данной подписке ограниено исключительно учебным процессом. Да и свобода ограничена некоммерческим использованием. После завершения обучения студенты должны деинсталлировать программные прдукты Microsoft, полученне по подписке, и в случае необходимости их дальнейшего испльзования приобретать коммерческие лицензии. С точки зрения факультета нет смысла обучать программным продуктам, за использование которых потом придется платить фирмам, в которые устроятся выпускники.

Третья причина — это свобода исследования, модификации и распространения кода, принятая в научной среде. Эту свободу может обеспечить только свободное ПО, и некторые проекты с открытым исходным кодом. Лицензионное соглашение Micosoft, предлагаемое в рамках MSDNAA, предоставляет данные свободы достаточно ограниченно.

Итак перед нами встала задача легализации программного обеспечения в рамках факультета. Внимание привлекло свободное и открытое программное обеспечение. Но какими критериями пользоваться при выборе этого ПО, а в некоторых случаях может лучше использовать коммерческий или проприетарный продукт?

1. Проблема выбора программного обспечения

Все мы заложники своих привычек, знаний и требований к себе и другим. Часто это вызывает проблемы во взаимоотношении между людьми, когда один специалист пытается вольно или невольно навязать другому свое мировосприятие. Многие Linux-оиды и Unix-оиды тоже заложники своих привычек. Среди сторонников Windows тоже немало людей, которых можно назвать ярыми поклонниками этой системы. Это хорошо, именно такие энтузиасты, способны поддерживать, развивать и популяризировать любимую операционную систеу. Но, часто их любовь перерастает в культ. А культ определенного дистрибутива, будь то Linux или MS Windows – это приводит к предвзятому отношению и необъективным оценкам работы в альтернативной системе.

Истинная свобода заключается НЕ в том, чтобы выбрать одну опреационную систему (ОС) навек: свободную или проприетарную, с открытым исходным кодом или закрытую, бесплатную или платную. Истинная свобода заключается в том, что мы должны уметь выбрать подходящую ОС, исходя из решаемых нами задач в соответствии с определенным критерием.

Каков же критерий? У каждого он свой. Перечислим некоторые из них:

- Я знаю эту ОС хорошо, следовательно, буду решать поставленную задачу с помощью имеющихся знаний о ней.
- Эта ОС работает с любым типом оборудования, поэтому решение будет строиться на ней.
- Эта система обеспечивает наилучший уровень безопасности, ведь это критично в нашей системе.
- Эта ОС универсальна, и если что, на ее основе можно построить другое решение, когда текущая задача будет решена.
- Я принимаю только решения на основе бесплатных или открытых решений.
- Я выбираю наиболее эффективную с точки зрения функционирования систему для данной задачи.
- Я выбираю наиболее экономически эффективную систему по внедрению и эксплуатации.

Можно продолжать, уточняя критерии. Из всего выделить главное: композитный критерий, включающий три компоненты:

- Время внедрения.
- Стоимость внедрения и эксплуатации.
- Эффективность (обычно, быстродействие и надежность) работы системы в рамках решаемых задач.

Исходя из задач, могут быть добавлены и дополнительные критерии такие, как безопасность, удобство управления системой и т.п. Но указанные три критерия будут присутствовать в любом проекте.

Например, для создания сетевой инфраструктуры скорее всего воспользуемся оборудованием и операционной системой Cisco, потому что именно она заточена для решения сетевых задач. Её эффективность доказана многими проектными решениями. Правда стоимость ее внедрения может остановить, если масштаб задачи не слишком велик, тогда её экономическая эффективность не будет оправдана.

Очевидно, что нельзя выбрать одну ОС для решения всех задач. Каждая ОС имеет свои преимущества и недостатки. Чтобы получить свободу выбора мы должны в целом представлять достоинства, недостатки каждой системы или хотя бы представителей семейства. Те же рассуждения применимы и к системам управления базами данных (СУБД), языкам программирования и другим программным продуктам.

У Microsoft и Sun Microsystems разные бизнес модели. Не нужно думать, что какие-то фирмы, исходя из своей доброй воли и альтруистических побуждений, делают использование своего программного обеспечения свободным и бесплатным (Sun, IBM). Если они от этого не будут в перспективе получать прибыль (как правило, от продажи поддержки), то они просто разоряться и некому будет поддерживать это открытое и бесплатное ПО. Время энтузиастов из университетских городков, создающих такое ПО для своих научных нужд закончилось. Нет, конечно, такие люди остались, но их мало и они не в состоянии в одиночку вести борьбу за свободу и открытость ПО.

С точки зрения обучения нужно выбирать наиболее перспективные универсальные ОС, лидеров в своем направлении, с учетом заложенного в них потенциала развития. С другой стороны, нужно избежать повторений и слишком сильных пересечений в курсах по администрированию и программированию, особенно по программированию. Каковы цели вуза? Создать конкурентный продукт на рынке труда. Выпускник должен обладать такими знаниями, которые позволили бы ему быстро принять наиболее эффективное решение поставленных перед ним задач. Он должен обладать способностью к изучению новых перспективных технологий.

Подведем итоги, распротранив выводы с ОС на языки, среды программирования и СУБД:

1. Учить нужно разным ОС.

2. Этих ОС не должно быть много (не более трех). Предлагаемый список на настоящий момент, с которым можно спорить: MS Windows Server 2003/2008, OpenSolaris 2008.11/ 2009.05, CiscoIOS v.12. С точки зрения администрирования важно изучать именно серверных представителей и иметь пример специализированной ОС, например CiscoIOS.

3. Это должны быть наиболее интересные и перспективные в дальнейшем своем развитии ОС. Все три указанные системы обладают этими свойствами.

4. Учить нужно как администрированию этих ОС так и программированию под них, если это возможно (Cisco IOS закрытая проприетарная система, под неё программировать затруднительно). Причем как скриптовым (пакетным) языкам так и базовому языку, на котором написано ядро системы и драйверы. Это, как правило, язык C.

5. Программирование на языках высокого уровня необходимо изучать на межплатформенном универсальном языке. Выбор языка под задачу важен, но наша задача – обучение, а потому важна именно универсальность и многоплатформенность. Хорошо подходит для этого Java. C# и платформу .NET – придется отбросить в виду их явной направленности на ОС MS Windows. Существующие реализации .NET типа Mono на Linux будут постоянно отставать в развитии от генеральной линии Microsoft.

   Есть, конечно, масса интересных языков это и JRuby, и lo, и много других замечательных языков. Но пока по статистике Java является самым популярным языком программирования [2], большая часть корпоративных приложений пишется на Java, переубедить отказаться от Java будет трудно.

SQL как язык взаимодействия с базами данных необходимо изучать там, где идет речь о СУБД.

Обязательно включить в программу курса (или отдельным курсом) изучение технологий разработки и шаблонов проектирования от GRASP и GoF [3] и далее с переключением на шаблоны проектирования для корпоративных приложений, [4]. Конечно, обязателен к изучению рефакторинг [5] и методика тестирования с JUnit или CUnit в зависимсти от изучаемого языка.

6. Из универсальных скриптовых языков столило изучать: 1) PHP, как наиболее распространенный язык для web-программирования; 2) Python – по-мнению многих экспертов, один из мощнейших интерпритируемых языков, имеющий широкое распространение. Кроме того, оба имеют хорошую связку с базами данных.

7. Учить нужно и разным СУБД и их также не должно быть много: 1) Oracle, как лидер отрасли; 2) MySQL, как наиболее используемая для построения web-контента, универсальная и бесплатная. Ну может из коммерческих MS SQL, но не больше. А вообще должен быть еще и академический курс по базам данных, непривязанный к конкретным СУБД.

8. Инструменты разработчика зависят от ОС. Но, к счастью не всегда. NetBeans – многоплатформенная IDE с поддержкой: Java, C, C++, PHP, Ruby и JRuby. Альтернативой является Eclipse, но с точки зрения обучения Netbeans предпочтительнее, так как использует идеалогию «все в одном флаконе» и располагается в нескольких комплектациях в зависимости от ваших потребностей и возможностей.

9. Все предыдущие предметы должны быть включены в основную программу обучения. Все инновации следует преподавать в спецкурсах. Не все инновации приживаются, поэтому нет смысла их внедрять в основную программу.

10. Самое трудное подобрать замену математическим пакетам MatLab и Maple. К сожалению из бесплатных и свободных продуктов полных аналогов со всей функциональностью найти невозможно. Наиболее сильная замена для MatLab пожалуй Octave, а для Maple – Maxima

11. Издательские системы. Выбор очевидно падает на свободный TEX. Open Office должен установлен на всех компьютерах факультет, но освоить его можно и самостоятельно.

Эти выводы представляют собой программу, которая реализуется на факультете. В настоящий момент изучение ряда дисциплин специальностей «Компьютерная безопасность» и «Вычислительные машины, комплексы, системы и сети» ведется в соотвесвии с этой программой: преподаванние Java-технологий внесено в основную программу обучения, администрирование Cisco оборудования изучается уже течении пяти лет, на факультете преподаются учебные курсы Oracle. Примером успешного использования Java-технологии в предметах специализации является курс «Криптографические протоколы» [6].

2. Академические программы

После выбора ПО следует выбор академической программы производителя ПО. Самые известные академические программы:

- Cisco Network Academy
- Microsoft IT Academy и Microsoft Development Network Academic Alliance (MSDNAA)
- Sun Java Academy и Sun Campus Ambassador
- Oracle Academy
- IBM Academic Initiative
- Red Hat Academy
- Центр компетенции Mandriva Linux
- Kaspersky Lab Academy

Участие в академической программе, имеет ряд преимуществ: предоставляется бесплатное обучение для преподавателей, льготное (для коммерческого ПО) и / или бесплатное предоставление продуктов для использования в учебном процессе и исследовательской работе, поддержка семинарами, льготная продажа оборудования, предоставление вакансий для выпускников и т.д.

Отметим, что на факультете уже работают академические программы Cisco, Microsoft, Sun, Oracle, Mandriva. В трех программах используется проприетарное, коммерческое программное ПО: Cisco, Microsoft и Oracle. В двух – откытое с различными полусвободными и свободными лицензиями: Sun и Mandriva.

Выбор академической программы в общем случае также сложен как и выбор программного обеспечения и может осуществляться на основе метода анализа иерархий [7].

3. Трудности

Перечислим основные трудности с которыми приходиться сталкиваться при переходе на новое ПО

- Издательские системы.
  - Требования научных издательств присылать статьи в формате Word.
  - Не возможность перехода на TEXв системе документооборота вуза.
  - Наличие заготовок и шаблонов документов в формате Word-a.
  - Не смотря на то, что Open Office читает все форматы Microsoft Office полного соответствия в форматировании не наблюдаетя. Особенно

это проявляется в математических формулах. Это приводит к тому, что приходиться перерабатывать разаработанные ранее методические указания и лабораторые работы.

- Операционные системы, прикладное и инструментальное ПО. Серьезных технических проблем не возникает, все проблемы решаются с помощью MSDNAA-подписки и систем виртуализации, например использования VirtualBox. Существуют проблемы психологического характера у преподавателей.
  - Необходимость перерабатывать материалы курсов с учетом специфики нового ПО.
  - Страх перед незнакомыми техникой и программным обеспечением.
  - Сложившиеся психологические установки: ошибки в работе хорошо знакомого прогаммного обеспечения расцениваются менее критично

  Обучение и переобучение может устранить переисленные проблемы. Важно, чтобы стоимость обучения и переобучения не превышали затрат на эксплуатацию коммерческого ПО.

Заключение

Выбор ПО для учебного процесса непростая задача. С одной стороны необходимо обеспечить разнообразие ПО с другой стороны его не должно быть слишком много, так как множественный выбор снижает глубину получаемых знаний. Нужна некоторая разумная достаточность и как кажется автору данной статьи этот балланс достигнут в настоящее время на ФКН.

По возможности нужно выбирать свободное ПО или хотябы ПО с открытым исходным кодом. При отсутствии необходимого свободного или открытого ПО в рамках решаемых учебных задач, выбор падает на коммерческое ПО. Производители, лидеры своей отрасли предлагают академические программы, участие в которых позволяет минимизировать расходы по приобретению ПО и обучению и переобучению преподавателей.

Постоянно быть в курсе новых разработок, технологий, инструментов и оборудования можно только при тесном взаимодействии с производителями, что обеспечивается совместным участем в академических программах.

Литература

1. Столлмен Р. Почему «открытый код» проигрывает свободному программному обеспечению.
   URL: `http://www.gnu.org/philosophy/open-source-misses-the-point.ru.html`
2. TIOBE Programming Community Index for May 2009.
   URL: `http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html`
3. Ларман К. Применение UML и шаблонов проектирования. Издательство: Вильямс, 2002. 624 с.
4. Фаулер М. Архитектура корпоративных программных приложений. М.: «Вильямс», 2007. С. 544.

5. Фаулер М. Рефакторинг. Улучшение существующего кода. Издательство: Символ-Плюс, 2008. 432 с.
6. Богаченко Н.Ф., Лавров Д.Н. Использование языка Java и среды разработки NetBeans на занятиях по дисциплине «Криптографические протоколы» // Открытые информационные технологии: пути развития и внедрения: Материалы российской научно-практической конференции. - Уфа: Восточный университет,2008. С.65-69.
7. Богаченко Н.Ф., Лавров Д.Н. Применение метода анализа иерархий в задаче выбора академической программы. Научно-технические ведомости СпбГПУ. Информатика. Телекоммуникации. Управление. Санкт-Петербург. Издательство Политехнического университета, 2008. № 3(60). С.187-191

Математические структуры
и
моделирование

Выпуск 19

Журнал

Редактор Е.В. Брусницына

Подписано в печать 20.09.2009.
ОП. Формат 60 × 84 1/8. Печ.л. 12,7. Уч.-изд.л. 21,0.
Тираж 100 экз.

Отпечатано в Полиграфическом центре КАН
644050, г. Омск, пр. Мира, 32, ком. 11, тел. (381-2) 65-47-31
Лицензия ПЛД № 58-47 от 21.04.97 г.