

М. А. Райфельд, А. А. Спектор

(Новосибирск)

НЕПАРАМЕТРИЧЕСКИЙ МЕТОД ОБНАРУЖЕНИЯ СИГНАЛОВ ОТ СЕЙСМИЧЕСКИ АКТИВНЫХ ОБЪЕКТОВ

Рассматривается обнаружение сигналов в сейсмической системе наблюдения, возникающих, в частности, при движении человека по поверхности грунта. Вводятся математическая модель сигнала в виде периодической последовательности импульсов с неизвестным периодом повторения и со случайными значениями и модель сейсмического фона, имеющая характер непрерывного гауссовского процесса с неизвестной спектральной плотностью мощности. Исследуются адаптивные байесовская и непараметрическая процедуры обнаружения, их цифровые реализации. Дан анализ характеристик обнаружения.

Введение. Одним из способов контроля перемещений различных объектов по поверхности земли является использование сейсмических колебаний, возбуждаемых при движении [1–3]. Первичные датчики с относительно небольшим радиусом действия объединяются в антенны, в совокупности покрывающие контролируемые участки. Возникающие при возбуждении поверхности сигналы передаются от всех датчиков в общий электронный блок, где происходит их преобразование в цифровую форму перед последующей обработкой. В результате текущего компьютерного анализа образуется полная картина обстановки на участке наблюдения.

Сейсмические системы наблюдения (ССН) относятся к числу систем постоянного действия, поэтому их работа предполагает полную автоматизацию процессов извлечения полезной информации. Условия использования ССН часто характеризуются значительными по уровню и разнообразными по происхождению сейсмическими помехами. Поэтому, как в радиолокации и радионавигации, извлечение информации здесь включает решение задач обнаружения сигналов, их классификации, оценивания параметров.

Важнейшим эксплуатационным требованием к ССН является обнаружение без задержки объекта наблюдения в момент его появления в рабочей зоне системы. Учитывая, что отдельные сейсмодатчики размещаются обычно на максимально приемлемых расстояниях друг от друга (при которых в рабочей зоне отсутствуют ненаблюдаемые участки, вызванные ограниченной чувствительностью датчиков), первичное обнаружение необходимо осуществлять по их сигналам. Это дает возможность получателю информации реагировать на ее поступление в реальном времени. Обнаружение траектории объекта и определение его местоположения выполняются позже и с задержкой. Эти за-

дачи решаются путем совместной обработки сигналов группы датчиков, в окрестности которых происходит перемещение объекта.

Первичное обнаружение предъявляет требование стабилизации вероятности ложного обнаружения, поскольку любое обнаружение, в том числе и ложное, влечет за собой реакцию наблюдателя, обычно сопровождаемую значительными расходами, включая и материальные. Это означает, что первичный обнаружитель должен отвечать критерию Неймана – Пирсона, а при его разработке следует использовать процедуры, обладающие устойчивостью по отношению к меняющимся характеристикам сейсмического фона.

В работах последнего времени [1–6] по анализу сейсмических сигналов перечисленным обстоятельствам внимания уделяется недостаточно, в некоторых из них говорится о «гарантированном обнаружении». Перенесение акцента на создание автоматического обнаружителя, работающего в реальном времени, требует решения проблем первичного обнаружения, что и является предметом предлагаемой работы.

Математическая модель сейсмосигналов и их байесовское обнаружение. При ударном воздействии на поверхность земли в ней возбуждаются сейсмические волны, достигающие датчика по многим путям в поверхностных слоях грунта. Совместное воздействие грунта и датчика на каждую волну может быть представлено как узкополосная линейная фильтрация исходного возмущения, а наблюдаемый на выходе датчика сигнал – как результат сложения элементарных волн. Большое число слагаемых приводит к нормализации наблюдаемых сигналов, результатом чего является использование во многих работах гауссовских моделей [4–6]. Существенное отличие области сигнала от области помехи состоит в различии мощностей: в области сигнала наблюдаемые данные имеют мощность, равную сумме мощностей полезного сигнала и помехи.

При воздействии шагов человека соответствующие сигналы имеют характер некогерентных импульсов со случайными амплитудой и фазой. Помеха представляет собой непрерывный гауссовский процесс с неизвестным видом функции спектральной плотности мощности. Первым элементом обработки наблюдаемых сигналов в этих условиях является их адаптивное выбеливание, при котором отсчеты входного сигнала преобразуются в дельта-коррелированные гауссовские последовательности. Пусть $x_i, i=1, 2, \dots,$ – отсчеты исходного (окрашенного) процесса, имеющего корреляционную функцию $B_x(\tau)$. Воспользуемся моделью марковского процесса порядка J для описания этого процесса, согласно которой [6]

$$x_i = \sum_{j=1}^J a_j x_{i-j} + \xi_j. \quad (1)$$

Здесь $\mathbf{A} = \|a_1, a_2, \dots, a_J\|^T$ – вектор параметров (Т – транспонирование), а ξ_j – информационный белый шум (ИБШ). Сумма в правой части описывает механизм предсказания текущего отсчета x_i на основе J предыдущих отсчетов, а ИБШ ξ_j интерпретируется как ошибка предсказания. Минимизируя ее дисперсию, нетрудно прийти к системе линейных алгебраических уравнений вида $\mathbf{B}_n \mathbf{A} = \mathbf{B}_n$, решением которой для вектора \mathbf{A} будет

$$\mathbf{A} = \mathbf{B}_n^{-1} \mathbf{B}_n.$$

Здесь матрица $\mathbf{B}_л$ левой части и вектор $\mathbf{B}_п$ правой составлены из значений корреляционной функции $B_x(\tau)$:

$$\mathbf{B}_л = \begin{vmatrix} B_x(0) & B_x(1) & \dots & B_x(J-1) \\ B_x(1) & B_x(0) & \dots & B_x(J-2) \\ \dots & \dots & \dots & \vdots \\ B_x(J-1) & B_x(J-2) & \dots & B_x(0) \end{vmatrix},$$

$$\mathbf{B}_п = \| B_x(1) \ B_x(2) \ \dots \ B_x(J) \|^T.$$

При известных коэффициентах выбеливание выполняется, как следует из (1), при помощи процедуры

$$\xi_j = x_i - \sum_{j=1}^J a_j x_{i-j},$$

применение которой предполагает предварительное оценивание корреляционной функции фона $B_x(\tau)$ для определения вектора параметров предсказания \mathbf{A} . Эксперименты показывают, что в условиях крупного производства с высоким уровнем сейсмического фона эффективное выбеливание достигается при $J = 8, \dots, 12$.

Представим полученные после выбеливания данные на входе обнаружителя в виде

$$u_j = \begin{cases} s_j \xi_j, & \text{если сигнал есть;} \\ \sqrt{d_0} \xi_j, & \text{если сигнала нет,} \end{cases} \quad j \in [0, (n-1)T + m - 1],$$

где сигнальная функция s_j описывает периодическую последовательность n импульсов, имеющих длительности m и период T следования отсчетов:

$$s_j = \begin{cases} \sqrt{d_1} & \text{при } j \in [(k-1)T, (k-1)T + m - 1], \quad k = \overline{1, n}; \\ \sqrt{d_0} & \text{при других значениях } j. \end{cases} \quad (2)$$

Согласно этой модели полезный сигнал имеет вид последовательности n периодически следующих пакетов, содержащих по m независимых отсчетов с гауссовским распределением и дисперсией d_1 . Промежутки между отсчетами сигнала заполнены отсчетами помехи, дисперсия которых $d_0 < d_1$.

Обнаружитель Неймана – Пирсона при известном значении периода T следования импульсов вычисляет отношение правдоподобия (или какую-либо его монотонную функцию) наблюдаемой выборки и сравнивает ее с порогом. В данном случае достаточной статистикой является сумма квадратов

наблюдаемого сигнала (поэтому обнаружитель называют энергетическим приемником), и правило принятия решения имеет вид

$$z(\mathbf{U}) = \sum_{j \in \Omega(T)} u_j^2 \begin{cases} \geq z_0 \Rightarrow \text{сигнал есть;} \\ < z_0 \Rightarrow \text{сигнала нет.} \end{cases} \quad (3)$$

Область суммирования $\Omega(T)$ в (3) включает точки j , соответствующие участкам предполагаемого наличия сигнальных элементов (первая строка формулы (2)), а вектор \mathbf{U} – данные, наблюдаемые в точках множества $\Omega(T)$. Порог z_0 определяется требуемым значением вероятности ложной тревоги F .

При обнаружении сейсмических сигналов, вызванных шагами человека, период T следования импульсов неизвестен, однако можно указать диапазон значений $[T_{\min}, T_{\max}]$, в пределах которого находится его значение. Для обнаружения сигнала в этих условиях используем метод обобщенного правдоподобия [7], при котором решающая статистика определяется выражением

$$L(\mathbf{U}) = \frac{\max_{T \in [T_{\min}, T_{\max}]} w_1(\mathbf{U}_T)}{w_0(\mathbf{U}_T)},$$

где \mathbf{U}_T – вектор входных данных, соответствующих сигнальной области $\Omega(T)$. Использование этого принципа приводит к решающему правилу вида

$$\begin{cases} z(\mathbf{U}) = \max_{T \in [T_{\min}, T_{\max}]} z_T(\mathbf{U}) \begin{cases} \geq z_0 \Rightarrow \text{сигнал есть;} \\ < z_0 \Rightarrow \text{сигнала нет;} \end{cases} \\ z_T(\mathbf{U}) = \sum_{j \in \Omega(T)} u_j^2. \end{cases} \quad (4)$$

Статистика $z(\mathbf{U})$, определяемая (3), подчиняется χ^2 -распределению с плотностью [8]:

$$w_{0(1)}(z) = \Gamma^{-1}(M/2) 2^{-M/2} d_{0(1)}^{-1} (z/d_{0(1)})^{(M/2)-1} e^{-z/2d_{0(1)}}. \quad (5)$$

Здесь индекс 0 означает отсутствие сигнала, а 1 – его наличие, число степеней свободы $M = mn$. Вероятности ложной тревоги F_T и правильного обнаружения D_T правила (3) при известном периоде T следования импульсов определяются выражениями:

$$F_T = 1 - \Gamma(z_0/2d_0, M/2); \quad D_T = 1 - \Gamma(z_0/2d_1, M/2), \quad (6)$$

в которых $\Gamma(x, a) = \Gamma^{-1}(a) \int_0^x e^{-t} t^{a-1} dt$ – функция χ^2 -распределения [9].

Полагая величины $z_T(\mathbf{U})$ при различных значениях T независимыми, для интегрального распределения максимальной статистики (4) имеем [10] $P(z) = P_T^{n_T}(z)$, где $P_T(z)$ – интегральный закон распределения величины $z_T(\mathbf{U})$, а n_T – число возможных значений периода повторения. Отсюда вероятность ложной тревоги правила (4)

$$F = 1 - (1 - F_T)^{n_T} \approx n_T F_T. \quad (7)$$

Вероятность правильного обнаружения D находим в предположении, что сигнальные элементы присутствуют лишь при одном значении периода T , откуда

$$D = 1 - (1 - F_T)^{n_T - 1} (1 - D_T) \approx D_T. \quad (8)$$

Приближенные равенства в (7) и (8) справедливы при $F_T \ll 1$.

Непараметрическое обнаружение сейсмосигналов. Основным недостатком байесовского обнаружения является необходимость точного знания дисперсии фона d_0 , поскольку порог обнаружения z_0 , выбираемый по заданной вероятности ложной тревоги F , определяется ее значением. Возможно применение адаптивного подхода, при котором находится оценка дисперсии d_0 , используемая затем для определения порога. В данном случае, однако, приходится учитывать, что реальные сейсмические сигналы фона содержат случайные неоднородности, среди которых наблюдаются импульсные увеличения интенсивности. Это препятствует применению классических оценок дисперсии, поскольку даже кратковременные увеличения интенсивности приводят к существенным смещениям получаемых оценок дисперсии и, следовательно, необходимого порога обнаружения.

Здесь предлагается иной подход, основанный на непараметрических статистиках. Рассмотрим двухвыборочную процедуру [11], в которой используются в качестве опорной выборка фона $\mathbf{U} = \|u_1, u_2, \dots, u_R\|^T$ и рабочая выборка $\mathbf{Y} = \|y_1, y_2, \dots, y_M\|^T$. Рабочая выборка получена в точках множества $\Omega(T)$ и поэтому соответствует гипотетическому значению периода $T \in [T_{\min}, T_{\max}]$. Элементы обеих выборок подчиняются распределению (5), причем параметром распределения для величин u_i является неизвестное значение d_0 , а для величин y_i – либо d_0 (если сигнала нет), либо d_1 (если сигнал есть).

Находятся минимальный и максимальный элементы опорной выборки:

$$u_{\min} = \min_i u_i, \quad u_{\max} = \max_i u_i,$$

и определяется число k_T элементов y_i рабочей выборки таких, что

$$y_i > u_{\max} \quad \text{или} \quad y_i < u_{\min}. \quad (9)$$

Аналогичные вычисления выполняются для всех n_T значений периода повторения импульсов, после чего находится решающая статистика

$$k = \max_{T \in [T_{\min}, T_{\max}]} k_T$$

и принимается решение в соответствии с правилом

$$k \begin{cases} > k_0 \Rightarrow \text{сигнал есть;} \\ \leq k_0 \Rightarrow \text{сигнала нет.} \end{cases}$$

Условная вероятность получить в рабочей выборке ровно k элементов, отвечающих неравенствам (9), при данном векторе \mathbf{U} определяется выражением

$$P(k | \mathbf{U}) = \frac{M!}{k!(M-k)!} [1 - F_y(u_{\max}) + F_y(u_{\min})]^k [F_y(u_{\max}) - F_y(u_{\min})]^{M-k}, \quad (10)$$

где $F_y(\cdot)$ – функция распределения элементов рабочей выборки \mathbf{Y} . Безусловная вероятность находится усреднением (10) по u_{\min} и u_{\max} :

$$P(k) = \iint P(k | \mathbf{U}) w_0(u_{\min}, u_{\max}) du_{\min} du_{\max}, \quad (11)$$

где плотность совместного распределения минимальной и максимальной порядковых статистик $w_0(u_{\min}, u_{\max})$ определяется выражением [10]

$$\begin{cases} w_0(u_{\min}, u_{\max}) = \frac{R!}{(R-2)!} w_0(u_{\min}) w_0(u_{\max}) [F_0(u_{\max}) - F_0(u_{\min})]^{R-2}; \\ u_{\min} < u_{\max}. \end{cases}$$

Вероятность решения в пользу наличия сигнала

$$P(k > k_0) = \sum_{k=k_0+1}^M P(k) \quad (12)$$

является локальной вероятностью ложной тревоги F_T при подстановке в (10) $F_y(\cdot) = F_0(\cdot)$ и локальной вероятностью правильного обнаружения D_T при $F_y(\cdot) = F_1(\cdot)$, где $F_0(\cdot)$ и $F_1(\cdot)$ – функции распределения, соответствующие плотностям $w_0(\cdot)$ и $w_1(\cdot)$. Общие вероятности ложной тревоги F и правильного обнаружения D находятся затем из (7), (8).

При больших значениях M вычисления по формулам (10)–(12) с использованием численного интегрирования в (11) труднореализуемы, поэтому целесообразно воспользоваться асимптотическим приближением Муавра – Лапласа для биномиального распределения (10). В результате из (11) нетрудно получить

$$P(k > k_0) = \iint w_0(u_{\min}, u_{\max}) \left\{ \Phi \left[\frac{M(1-p)}{\sqrt{Mp(1-p)}} \right] - \Phi \left[\frac{k_0 + 1 - Mp}{\sqrt{Mp(1-p)}} \right] \right\} du_{\min} du_{\max}.$$

Здесь $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$ – гауссовская функция распределения; $p = 1 - F_y(u_{\max}) + F_y(u_{\min})$.

Для оценки потерь, связанных с применением описанных непараметрических статистик, на рис. 1, *a* приведены характеристики обнаружения одиночного импульсного сигнала (при $n = 1$). Сплошные линии соответствуют байесовскому, а штриховые – непараметрическому обнаружителю. Проигрыш непараметрического обнаружителя по величине порогового отношения сигнал/шум $q^2 = d_1/d_0$, как следует из рисунка, составляет около 3 дБ.

Согласно (7), (8) характеристики рис. 1 можно использовать и для оценки обнаружителя импульсного пакета при неизвестном периоде повторения импульсов с учетом того, что характеристики, рассчитанные для определенных значений локальной вероятности ложной тревоги F_T , теперь будут справедливы для вероятности ложной тревоги F , увеличенной в соответствии с (7) в n_T раз. Из этого, например, следует, что при $n_T = 10$ характеристики рис. 1 можно использовать, заменив в них значения $F_T = 10^{-\Theta}$ значениями $F = 10^{-\Theta+1}$. Энергетические потери, связанные с неопределенностью периода повторения импульсов, составляют всего 0,26 дБ для байесовского и 0,42 дБ для рассматриваемого непараметрического обнаружителей сигналов.

Влияние размера опорной выборки на характеристики обнаружения непараметрического обнаружителя иллюстрирует рис. 1, *b*. Здесь представлены зависимости порогового отношения сигнал/шум, обеспечивающего обнаружение сигнала с вероятностью 0,95, от объема опорной выборки R при нескольких значениях вероятности ложной тревоги и $n_T = 1$. Увеличение размера опорной выборки приводит к снижению требуемого энергетического потенциала. Особенно сильна эта зависимость при $R < M$. Увеличение же R при $R > M$ проявляется значительно слабее. Учитывая, что рост размера опорной выборки приводит к возрастанию объема вычислений, реализующих обнаружение, можно рекомендовать использование одинаковых размеров рабочей и опорной выборок.

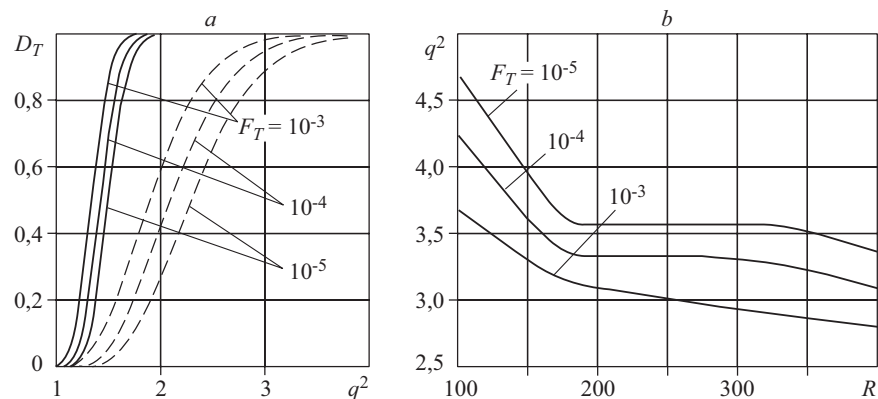


Рис. 1. Характеристики обнаружения байесовского и непараметрического обнаружителей ($M = 200, R = 400$): рабочие характеристики (*a*); влияние размера опорной выборки на характеристики обнаружения (*b*)

Цифровое обнаружение сейсмосигналов движущегося объекта. На практике удобно использовать двухэтапные обнаружители сигналов, у которых на втором этапе применяется цифровое накопление. На первом этапе обработки принимаются предварительные решения о наличии или отсутствии сигнала в каждом из предполагаемых импульсов обнаруживаемого сигнала. Получаемые бинарные данные поступают затем в счетчик единиц. Пусть x_{i+jT} – результат (0 или 1) первичной обработки гипотетического j -го импульса пачки, имеющей гипотетический период T следования импульсов. Тогда решение о наличии или отсутствии сигнала принимается в соответствии с правилом

$$v_i = \max_{T \in [T_{\min}, T_{\max}]} \{v_i(T)\} \begin{cases} > v_0 \Rightarrow \text{сигнал есть;} \\ \leq v_0 \Rightarrow \text{сигнала нет,} \end{cases}$$

где $v_i(T) = \sum_{j=0}^{n-1} x_{i+jT}$, а v_0 – цифровой порог. Пусть θ_0 – вероятность того, что случайная величина x_{i+jT} принимает значение 1 при фактическом отсутствии сигнала. Тогда согласно (7) для общей вероятности ложной тревоги справедливо выражение

$$F \approx n_T F_T = n_T \sum_{v=v_0+1}^n \binom{n}{v} \theta_0^v (1-\theta_0)^{n-v}.$$

Для вероятности правильного обнаружения справедлива согласно (8) формула

$$D \approx D_T = \sum_{v=v_0+1}^n \binom{n}{v} \theta_1^v (1-\theta_1)^{n-v},$$

в которой θ_1 – также вероятность события $x_{i+jT} = 1$, но при реальном наличии во входных данных полезного сигнала.

При использовании байесовского принципа первичная обработка заключается в накоплении квадратов наблюдаемых (выбеленных) отсчетов, но в отличие от (3) суммирование выполняется отдельно для каждого ожидаемого импульса. Получаемые при этом величины z_{i+jT} , $j=0, n-1$, сравниваются с аналоговым порогом z_0 , при превышении которого $x_{i+jT} = 1$, а в противном случае $x_{i+jT} = 0$. Очевидно, что в этом случае вероятность θ_0 определяется первым из соотношений в (6), θ_1 – вторым, причем в этих соотношениях теперь число степеней свободы вместо M должно быть взято равным m – длительности одного импульса.

При использовании непараметрического принципа первичной обработки в роли рабочей выборки для j -го импульса $\mathbf{Y}_j = \|y_{j,1}, y_{j,2}, \dots, y_{j,m}\|^T$ выступают m его отсчетов, а в роли опорной $\mathbf{U}_j = \|u_{j,1}, u_{j,2}, \dots, u_{j,r}\|^T$ какие-либо r соседних отсчетов фона. Далее для этих выборок реализуется про-

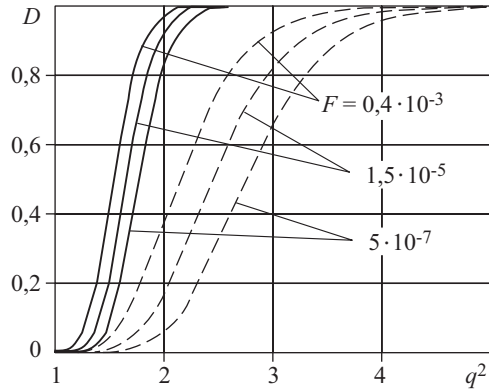


Рис. 2. Рабочие характеристики цифрового обнаружителя ($m = 50$, $r = 100$, $n = 4$, $n_T = 10$)

цедура подсчета числа элементов k_j вектора \mathbf{Y}_j , вышедших за пределы интервала, ограниченного минимальным и максимальным элементами вектора \mathbf{U}_j . Цифровые данные x_{jT} образуются путем бинаризации полученных чисел:

$$x_{jT} = \begin{cases} 1 & \text{при } k_j > k_0; \\ 0 & \text{при } k_j \leq k_0. \end{cases}$$

Вероятности θ_0 и θ_1 находятся теперь при помощи соотношения (12) подстановкой $F_y(\cdot) = F_0(\cdot)$ и $F_y(\cdot) = F_1(\cdot)$ соответственно. Кроме того, полные объемы опорной R и рабочей M выборок заменяются значениями r и m , соответствующими одному импульсу полезного сигнала.

На рис. 2 приведены характеристики описанного цифрового обнаружителя. Сплошные кривые относятся к байесовской первичной обработке, штриховые – к непараметрической. Общие объемы выборок $M = mn$, $R = rn$ на этом рисунке совпадают с аналогичными параметрами рис. 1. Применение цифрового накопления приводит к небольшим потерям, составляющим по величине порогового отношения сигнал/шум 0,7 дБ для байесовской процедуры первичного накопления и 0,3 дБ для непараметрической обработки. При этом несколько сокращается величина потерь непараметрической процедуры по отношению к байесовской, которая теперь составляет 2,6 дБ.

Заключение. Предложенные в работе методы обнаружения сейсмических сигналов, возникающих при движении человека, с несущественной модификацией можно использовать для обнаружения других сейсмически активных объектов: автомобилей, крупных животных и др. Натурные эксперименты подтверждают полученные в работе теоретические данные. Предложенный непараметрический метод обладает устойчивостью своих характеристик, что является чрезвычайно важной его чертой. Вместе с тем свойственные ему заметные потери эффективности (в сравнении с байесовским обнаружителем) свидетельствуют о необходимости дальнейшего поиска более эффективных процедур обнаружения сигналов.

СПИСОК ЛИТЕРАТУРЫ

1. **Введенский Б. С.** Периметровые охранные системы фирмы “Geogip” // Системы безопасности связи и телекоммуникаций. 1999. № 27.
2. **Иванов В. А., Онуфриев Н. В.** Развитие принципов адаптации сейсмических средств охраны участков местности // Радиотехника. 2005. № 3. С. 97.
3. **Крюков И. Н.** Математическая модель подсистемы обнаружения сейсмических средств обнаружения территориально распределенных радиотехнических систем охраны // Там же. С. 84.
4. **Воскобойникова Г. М.** Повышение помехоустойчивости выделения и измерение параметров вибрационных сейсмограмм методом динамического программирования // Мат. междунар. техн. конф. «Информационные системы и технологии». Новосибирск: НГТУ, 2003. С. 95.
5. **Хайретдинов М. С., Омельченко О. К., Родионов Ю. И.** Проблема автоматизированного восстановления сейсмического источника // Там же. С. 176.
6. **Анализ** и выделение сейсмических сигналов. М.: Мир, 1986.
7. **Левин Б. Р.** Теоретические основы статистической радиотехники. М.: Сов. радио, 1975. Кн. 2.
8. **Королюк В. С., Портенко В. И., Скороход А. В., Турбин А. Ф.** Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.
9. **Большев Л. Н., Смирнов Н. В.** Таблицы математической статистики. М.: Наука, 1983.
10. **Дейвид Г.** Порядковые статистики. М.: Наука, 1979.
11. **Райфельд М. А.** Непараметрический алгоритм различения стохастических сигнала и помехи, отличающихся дисперсиями // Изв. вузов. Сер. Радиоэлектроника. 1991. № 1. С. 25.

*Новосибирский государственный
технический университет,
E-mail: spector@ktor.ref.nstu.ru
ktor@ref.nstu.ru*

*Поступила в редакцию
15 февраля 2005 г.*