СТАТИСТИЧЕСКИЙ АНАЛИЗ СОЦИАЛЬНО-ПСИХОЛОГИЧЕСКИХ ФАКТОРОВ УВЕЛИЧЕНИЯ ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ ЧЕЛОВЕКА

***К.А. ОСИПЕНКО*, Н.Б.ОСИПЕНКО**

The testing methodology provides the correct analysis of the data using different types of signs was described. As an example, select the task to identify the factors of human life extension

Ключевые слова: методика, апробация, корреляционный, регрессионный и дискриминантный анализы данных, разнотипные признаки

Работа посвящена описанию разработанной и апробированной на практическом примере [1] методики, обеспечивающей корректность регрессионного анализа данных при использовании разнотипных признаков. Настоящая методика позволяет решить следующие проблемы регрессионных построений: 1) неоднородность исходной выборки (за счет экспертного разбиения исходной выборки на подвыборки с использованием базовых классификационных признаков); 2) необеспеченность монотонности изменения целевого количественного показателя с ростом значений объясняющего порядкового или количественного признака (осуществляется разбиение значений такого признака на несколько градаций с использованием корреляционного анализа его связей с целевым свойством); 3) для разных интервалов изменения целевого количественного признака характерны различные механизмы участия объясняющих факторов в процессах формирования целевого свойства (решением этой проблемы является построение регрессионной модели по каждому диапазону целевого свойства); 4) имеющиеся в распоряжении исследователя признаки не полностью описывают весь механизм формирования целевого свойства (в этом случае целью моделирования является не точная оценка прогнозного показателя, а только изучение характера влияния объясняющих признаков на формирование целевого свойства).

Апробация методики осуществлялась также на примере задачи распознавания групп риска смертности людей (в первую группу вошли умершие от сердечно-сосудистых заболеваний и несчастных случаев, во вторую – умершие от онкозаболеваний и других хронических болезней) по паспортным данным. В рамках пакета «Statistica» была написана программа обучения распознаванию групп риска. Перевод даты рождения в набор числовых качественных признаков осуществлен с помощью общеизвестного алгоритма Пифагора и описан в работе [2]. Для перевода имени и фамилии в признаки для распознавания групп риска использована числовая азбука [3]. Ошибка распознавания первого рода (отнесение объекта первого класса ко второму) составила 10/55, ошибка второго рода – 1/33. Доля отказов в распознавании – 39/88. Как видим, несмотря на небольшой объем выборки, можно утверждать, что паспортные данные вполне пригодны для включения их в список признаков при экспресс-диагностике второй группы риска на первом этапе скрининга здоровья населения.

Литература

1. *Осипенко, К.А.* Пример «выращивания» регрессионной модели социального явления на базе критерия правдоподобности ее интерпретации / К.А.Осипенко, Н.Б.Осипенко, А.Н. Осипенко // Проблемы физики, математики и техники. – 2013. – №4(17). –С.85-88.

2. *Осипенко, К.А.* Метод регрессионного моделирования продолжительности жизни по дате рождения / *К.А.Осипенко, Н.Б.Осипенко* // Творчество молодых 2012: сборник научных работ студентов и аспирантов УО «ГГУ им. Ф. Скорины»: в 2 ч. / Гомельский гос. ун-т им. Ф.Скорины; отв. ред. О.М. Демиденко. – Гомель, 2012. –Ч. 1. –С.194-197.

3. *Хигир, Б.Ю.* Число имени / Б.Ю. Хигир. –СПб.: Астрель, 2008. –42с.

РАСПОЗНАВАНИЕ И ОЦЕНИВАНИЕ ПАРАМЕТРОВ КРИПТОГРАФИЧЕСКИХ ГЕНЕРАТОРОВ ПО ВЫХОДНЫМ ПОСЛЕДОВАТЕЛЬНОСТЯМ

***В.Ю. ПАЛУХА*, Ю.С. ХАРИН**

The construction of informative features for the statistical recognition of pseudorandom number generators is considered. The approach to construction of features and the descriptions of the constructed features based on entropy, ranks and determinants of matrices are given. The nonlinear feedback shift register is approximated by the Markov chain with partial connections. The parameters of Markov chain and shift register are estimated

Ключевые слова: криптографические генераторы псевдослучайных последовательностей, статистическое распознавание, информативные признаки, статистическое оценивание параметров, малопараметрические модели цепей Маркова

Введение

Современные средства криптографической защиты информации (СКЗИ) используют случайные и псевдослучайные последовательности $x_1, x_2, \ldots \in V = \{0, 1\}$, которые должны быть близки по своим свойствам к равномерно распределённой случайной последовательности (РРСП) [1]. Гипотезу о том, что выходная последовательность генератора $\{x_t\}$ является равномерно распределённой, будем обозначать $H_* = \{\{x_t\}$ есть РРСП$\}$.

При проведении испытаний СКЗИ с целью оценки их надёжности возникает ряд задач, обозначенных в [1] как S1, S2, S3. Математической сущностью этих задач криптоанализа является статистическое распознавание генераторов случайных и псевдослучайных последовательностей, т.е. отнесение (классификация) наблюдаемой выходной последовательности генератора $x_1, x_2, \ldots, x_T \in V$ некоторой длины $T < +\infty$ к одному из L $(2 \le L < +\infty)$ классов $\Omega_1, \ldots, \Omega_L$. Множество классов определяется спецификой задачи. Наиболее значимым этапом решения задачи статистического распознавания является построение пространства M информативных признаков $\rho_1, \ldots, \rho_M$, несущих информацию о разделимости классов $\{\Omega_i\}$.

Для полного решения описанных выше криптоаналитических задач S1, S2, S3 необходимо не только определить тип генератора, но и оценить его параметры. Зачастую задача оценивания параметров математической модели генератора является трудоемкой. При рассмотрении элементов выходной последовательности как реализаций некоторой случайной величины на вероятностном пространстве естественной является аппроксимация детерминированной математической модели генератора вероятностной моделью. В теории вероятностей аналогом генераторов псевдослучайных последовательностей являются цепи Маркова высокого порядка. Существуют различные малопараметрические модели цепей Маркова, для которых известны алгоритмы оценивания их параметров.

В данной статье описывается построение информативных признаков для распознавания криптографических генераторов, описание их вероятностных свойств, а также аппроксимация криптографических генераторов малопараметрическими марковскими моделями.

1. Распознавание генераторов

1.1. Подход к построению информативных признаков

Будем предполагать, что выходная последовательность генератора $x_t \in V$ является случайной последовательностью на некотором вероятностном пространстве (Ω, F, P). Разобьём последовательность $X = x_1, x_2, \ldots, x_T$ на l фрагментов длины $s(n)$ $X_1^{(n)}, \ldots, X_l^{(n)}$. Пусть наблюдается некоторая статистика $a(n) = f(X, n)$ при различных параметрах длины фрагмента $n \in [n_-, n_+]$, $1 \le n_- < n_+$; $a_*(n) = E_{H_*}\{a(n)\}$ – математическое ожидание этой статистики при истинной гипотезе H_*. В качестве признака предлагается использовать уклонение $a(n)$ от математического ожидания в l_1 – метрике:

$$\rho = \frac{1}{n_+ - n_- + 1} \sum_{n=n_-}^{n_+} |a(n) - a_*(n)|. \qquad (1.1)$$

Чем ρ больше, тем больше свойства генератора отличаются от РРСП.

Признак (1.1) можно модифицировать, взяв нормированное уклонение:

$$\rho = \frac{1}{n_+ - n_- + 1} \sum_{n=n_-}^{n_+} \frac{|a(n) - a_*(n)|}{|a_*(n)|}. \qquad (1.2)$$

1.2 Распознавание с использованием признаков на основе рангов матриц

Разбиваем наблюдаемый ряд $x_1, x_2, \ldots$ на фрагменты длины $s(n) = n^2$ $X^{(1)}, X^{(2)}, \ldots \in V^{n^2}$: $X^{(k)} = (x_{(k-1)n^2+1}, \ldots, x_{kn^2})$. Используя k-й фрагмент $X^{(k)} = (x_1^{(k)}, \ldots, x_{n^2}^{(k)}) \in V^{n^2}$ выходной последовательности, построим $(n \times n)$ – матрицу

$$A^{(k)} = (a_{ij}^{(k)}) = \begin{pmatrix} x_1^{(k)} & \dots x_n^{(k)} \\ \dots & \dots \dots \\ x_{(n-1)n+1}^{(k)} & \dots x_{n^2}^{(k)} \end{pmatrix} \in V^{n\times n}. \tag{1.3}$$

Ранг этой матрицы отражает наличие функциональной зависимости в последовательности. Если $rank(A^{(k)}) = r < n$, то в матрице $A^{(k)}$ имеется $n - r$ линейно зависимых над V_n строк. Обозначим ранг матрицы (1.3) $r^{(k)} = rank(A^{(k)}) \in \{0, 1, ..., n\}$. Пусть наблюдается l фрагментов, т.е. $T = l \cdot n^2$. Определим статистику, имеющую смысл среднего относительного ранга:

$$v(n) = \frac{1}{nl}\sum_{k=1}^{l} r^{(k)}. \tag{1.4}$$

Теорема 1.1. При верной гипотезе H_* математическое ожидание статистики (1.4) равно

$$E_{H_*}\{v(n)\} = v_*(n) = \frac{1}{n}\sum_{j=0}^{n} q_{nj}\, j, \tag{1.5}$$

где $q_{nj} = P_{H_*}\{r^{(k)} = j\} = 2^{j(2n-j)-n^2}\prod_{i=0}^{j-1}\frac{(1-2^{i-n})^2}{1-2^{i-j}}, \quad j \in \{0,1,...,n\}$ [1].

На основании статистик $\{v(n) : n_- \le n \le n_+\}$ и их математических ожиданий (1.5) построим информативные признаки согласно (1.1):

$$\rho = \frac{1}{n_+ - n_- + 1}\sum_{n=n_-}^{n_+} |\, v(n) - v_*(n)\,|.$$

1.3 Распознавание с использованием признаков на основе определителей матриц

Определитель матрицы позволяет учитывать зависимости между элементами матрицы. Для наблюдаемой двоичной последовательности $x_1, x_2, ..., x_T \in V = \{0, 1\}$ вычислим следующие статистики, основанные на детерминантах $(n, T \in \mathbb{N}, \quad T \gg n^2)$:

$$\beta(1) = \frac{1}{T}\sum_{t=0}^{T-1} x_{t+1}^2, \quad \beta(2) = \frac{1}{l}\sum_{t=0}^{l-1}(D_t^{(2)})^2, D_t^{(2)} = \begin{vmatrix} x_{4t+1} & x_{4t+2} \\ x_{4t+3} & x_{4t+4} \end{vmatrix}, ...,$$

$$\beta(n) = \frac{1}{l}\sum_{t=0}^{l-1}(D_t^{(n)})^2, D_t^{(n)} = \begin{vmatrix} x_{n^2t+1} & x_{n^2t+2} & \dots & x_{n^2t+n} \\ x_{n^2t+n+1} & x_{n^2t+n+2} & \dots & x_{n^2t+2n} \\ \dots & \dots & \dots & \dots \\ x_{n^2t+(n-1)n+1} & x_{n^2t+(n-1)n+2} & \dots & x_{n^2t+n^2} \end{vmatrix}; \quad l = \left[\frac{T}{n^2}\right]. \tag{1.6}$$

Теорема 1.2. При верной гипотезе H_* математическое ожидание величины $(D_t^{(n)})^2$ равно

$$E_{H_*}\{(D_t^{(n)})^2\} = \beta_*(n) = \frac{(n+1)!}{4^n}, n = 2, 3, ..., t = 1, 2, \dots \tag{1.7}$$

Заметим, что при увеличении n значение $\beta_*(n)$ из (1.7) растёт экспоненциально. Поэтому на практике удобно использовать не $\beta(n)$, а $\ln\beta(n)$. На основании статистик (1.6) и их математических ожиданий (1.7) построим информативные признаки согласно (1.2):

$$\rho = \frac{1}{n_+ - n_- + 1}\sum_{n=n_-}^{n_+} \frac{|\ln\beta(n) - \ln\beta_*(n)|}{|\ln\beta_*(n)|}.$$

1.4 Распознавание с использованием признаков на основе энтропии

Пусть $p_{i_1,\dots,i_n} = P\{x_{t+1} = i_1, \dots, x_{t+n} = i_n\}$ – распределение вероятностей n-граммы $(x_{t+1},\dots,x_{t+n}) \in V_n$, которое предполагается не зависящим от $t \in \mathbb{N}$. Многомерная (n-мерная) энтропия Шеннона для фрагмента длины n равна [1]:

$$h(n) = -\sum_{i_1,\dots,i_n \in V} p_{i_1,\dots,i_n} \ln p_{i_1,\dots,i_n}. \tag{1.8}$$

Обозначим: $i = \sum_{j=1}^{n} 2^{j-1} i_j$ – представление числа $i \in \{0, 1, \dots, 2^n - 1\}$ в двоичной системе счисления, $p_i(n) = P\{\sum_{j=1}^{n} 2^{j-1} x_j = i\} = p_{i_1,\dots,i_n}$, $i = 0, \dots, 2^n - 1$. Пусть наблюдается l фрагментов $X^{(1)}$, …, $X^{(l)}$. Построим статистические оценки распределения вероятностей $\{p_i(n)\}$, $i = 0, \dots, 2^n - 1$:

$$\hat{p}_i(n) = \frac{1}{l}\sum_{k=1}^{l} \delta_{\overline{X}^{(k)},i}, \quad \overline{X}^{(k)} = \sum_{j=1}^{n} 2^{j-1} x_j^{(k)}, \quad \delta_{\overline{X}^{(k)},i} = \begin{cases} 1, \overline{X}^{(k)} = i; \\ 0, \overline{X}^{(k)} \neq i. \end{cases}$$

Используя подстановочный принцип, построим статистическую оценку энтропии (1.8):

$$\hat{h}(n) = -\sum_{i=0}^{2^n-1} \hat{p}_i(n) \ln \hat{p}_i(n). \tag{1.9}$$

В [2] предложено использовать приращение энтропии при увеличении длины рассматриваемого фрагмента. Пусть наблюдается последовательность $x_1, \dots, x_T \in V = \{0, 1\}$. Для удобства «зациклим» последовательность до длины $T + n - 1$: $x_{T+1} = x_1$, …, $x_{T+n-1} = x_{n-1}$.

Воспользуемся оценкой энтропии (1.9). Обозначим приращение оценки энтропии

$$g(T,n) = \hat{h}(n) - \hat{h}(n-1), \quad g(T,1) = \hat{h}(1). \tag{1.10}$$

Теорема 1.3. Если $T, n \to \infty$, $\frac{T}{2^n} \to \lambda > 0$ и верна гипотеза H_*, то $g(T, n)$ имеет асимптотически нормальное распределение с асимптотическим математическим ожиданием

$$E_{H_*}\{g(T,n)\} \sim g_*(\lambda) = e^{-\lambda}\sum_{k=1}^{\infty} \frac{\lambda^k \ln(k+1)}{k!}(e^{-\lambda}2^k - 1). \tag{1.11}$$

На основании статистик (1.10) и их математических ожиданий (1.11) построим информативные признаки согласно (1.1):

$$\rho = \frac{1}{n_+ - n_- + 1}\sum_{n=n_-}^{n_+} |g(T,n) - g_*(\lambda)|.$$

2. Статистическое оценивание параметров математических моделей генераторов

Большинство криптографических генераторов является рекуррентной функцией некоторого порядка. Вероятностной моделью таких функций является цепь Маркова высокого порядка. Успешное оценивание параметров цепи Маркова, аппроксимирующей генератор, позволит оценить параметры исходного генератора. В рамках исследования произведена аппроксимация регистра сдвига с нелинейной обратной связью цепью Маркова с частичными связями. Опишем математические модели цепи Маркова и регистра сдвига, после чего их связь станет очевидной.

Пусть на вероятностном пространстве (Ω, F, P) определена эргодическая цепь Маркова s-го порядка (ЦМ(s)) $x_t \in V = \{0, 1\}, t \in \mathbb{N}$:

$$P\{x_{t+1} = i_{t+1} \mid x_t = i_t, \dots, x_1 = i_1\} = P\{x_{t+1} = i_{t+1} \mid x_t = i_t, \dots, x_{t-s} = i_{t-s}\}, t > s. \tag{2.1}$$

Обозначим условное распределение вероятностей одношаговых переходов

$$p_{i_{s+1}|i_1,\ldots,i_s} = P\{x_{t+s} = i_{s+1} \mid x_t = i_1, \ldots, x_{t+s-1} = i_s\}, \quad \sum_{i_{s+1}} p_{i_{s+1}|i_1,\ldots,i_s} = 1, \quad i_1, \ldots, i_s, i_{s+1} \in V,$$

которое в силу однородности цепи Маркова не зависит от $t \in \mathbb{N}$.

Содержательный смысл Марковского условия (2.1) заключается в том, что распределение вероятностей элемента последовательности x_{t+1} зависит не от всей предыстории $x_1, \ldots, x_t$, а только от s предыдущих элементов $x_{t-s+1}, \ldots, x_t$. По такому принципу работает большинство криптографических генераторов.

На практике модель полносвязной цепи Маркова высокого порядка используется редко из-за экспоненциального роста числа параметров модели (2^s) с ростом порядка s. Вместо них используются малопараметрические модели. Одной из таких моделей является цепь Маркова s-го порядка с r частичными связями (ЦМ(s, r)), предложенная в 2004 г. [3]:

$$p_{i_{s+1}|i_1,\ldots,i_s} = q_{(i_{m(1)},\ldots,i_{m(r)}),i_{s+1}}, \quad i_1, \ldots, i_s, i_{s+1} \in V, M = \{m(1), \ldots, m(r)\}, I = \{1, \ldots, s\}, M \subseteq I.$$

Вероятности одношаговых переходов $q_{(i_{m(1)},\ldots,i_{m(r)}),i_{s+1}}, \quad i_1,\ldots,i_{s+1} \in V$, образуют матрицу вероятностей одношаговых переходов Q размерности $2^r \times 2$. Суть модели ЦМ(s, r) состоит в том, что следующий элемент последовательности x_{t+1} зависит не от всех s предыдущих элементов $x_{t-s+1}, \ldots, x_t$, а от некоторых $r \leq$ s элементов $x_{t-s+m_1}, \ldots, x_{t-s+m_r}$. Если $r = s$, то получаем полносвязную цепь Маркова, если $r < s$, то число параметров модели сокращается до 2^r. В этом заключается преимущество малопараметрических моделей цепей Маркова.

Одними из самых простых в реализации и удобных в использовании генераторов являются регистры сдвига с функциональной обратной связью. Эти генераторы характеризуются состоянием ($y_1\ y_2\ \ldots\ y_s$) и функцией обратной связи $f(y_1, y_2, \ldots, y_s)$, $y_i \in V = \{0, 1\}, i \in \{1, \ldots, s\}$. Приведем алгоритм выработки последовательности регистром сдвига с обратной связью.

- Вход: начальное состояние ($x_1\ x_2\ \ldots\ x_s$), длина последовательности T.
- Шаг 0: ($y_1\ y_2\ \ldots\ y_s$) := ($x_1\ x_2\ \ldots\ x_s$).
- Шаг i, $i = 1, \ldots, T$:
  - $x_i := y_1$; $c := f(y_1, y_2, \ldots, y_s)$.
  - Для $j = 1, \ldots, s - 1$: $y_j := y_{j+1}$.
  - $y_s := c$.
- Выход: $x_1, \ldots, x_T$.

Если функция f является линейной, то в этом случае генератор является регистром сдвига с линейной обратной связью (РСЛОС). РСЛОС является хорошо изученным с точки зрения криптоанализа генератором. Функция обратной связи восстанавливается при помощи алгоритма Берлекэмпа – Месси. Практический интерес представляют генераторы с нелинейной функцией – регистры сдвига с нелинейной обратной связью. Если функция f существенно зависит от r переменных $\{y_{m(1)}, \ldots, y_{m(r)}\}$, $\{m(1), \ldots, m(r)\} \subseteq \{1, \ldots, s\}$, то становится очевидной аналогия модели ЦМ(s, r) и математической модели регистра сдвига с обратной связью.

В [4] предложен алгоритм статистического оценивания параметров модели ЦМ(s, r). С помощью этого алгоритма строится оценка порядка цепи Маркова $s \in [s_-, s_+]$, числа частичных связей $r \in [r_-, r_+]$, шаблона M и матрицы Q. Вычислительная сложность алгоритма оценивания M и Q при известных r и s и длине последовательности T равна $O(2^{r+1}s^{r-1} + Ts^r)$.

Оценка матрицы вероятностей одношаговых переходов Q, построенная в результате работы алгоритма оценивания параметров ЦМ(s, r), может быть использована для построения таблицы истинности булевой функции обратной связи. Она строится следующим образом. Матрица Q имеет размерность $2^r \times 2$. Каждая строка матрицы $(q_{(i_{m(1)},\ldots,i_{m(r)}),0}\ q_{(i_{m(1)},\ldots,i_{m(r)}),1})$ содержит вероятности генерации 0 или 1 в зависимости от предыстории $i_{m(1)}, \ldots, i_{m(r)}$. Если вероятность генерации 0 больше вероятности генерации 1, то считаем, что булева функция принимает значение 0, в противном случае значение функции полагаем равным 1:

$$f(i_{m(1)},\ldots,i_{m(r)})=\begin{cases}0, & q_{(i_{m(1)},\ldots,i_{m(r)}),0} > q_{(i_{m(1)},\ldots,i_{m(r)}),1};\\ 1, & q_{(i_{m(1)},\ldots,i_{m(r)}),0} \le q_{(i_{m(1)},\ldots,i_{m(r)}),1}.\end{cases}$$

Заметим, что если мы имеем дело с генератором без искажений, то строки матрицы будут иметь вид (0 1) либо (1 0). Это значит, что вектором значений булевой функции является второй столбец матрицы Q.

Заключение

В данном исследовании решены следующие задачи:

1) Предложен общий подход к построению информативных признаков. Построены информативные признаки на основе энтропии, рангов и определителей матриц, вычислены математические ожидания используемых статистик. Т.к. предложенный подход не использует внутреннее строение генераторов, то построенные признаки можно применять для распознавания различных типов генераторов.

2) Предложен подход к статистическому оцениванию параметров генераторов при помощи аппроксимации их выходных последовательностей марковскими моделями. Данный подход успешно применён для оценивания параметров регистра сдвига с нелинейной обратной связью на основе цепи Маркова с частичными связями.

Литература

1. Криптология / Ю.С. Харин [и др.]. – Минск: БГУ, 2013. – 512 с.
2. *Rukhin, A.L.* Approximate Entropy for Testing Randomnesses / A.L. Rukhin. Режим доступа: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.6174&rep=rep1&type=pdf Дата доступа: 17.09.2014
3. *Харин, Ю.С.* Цепи Маркова с r частичными связями и их статистическое оценивание / Ю.С. Харин // Доклады НАН Беларуси, 2004. – Т. 48, № 1. – С. 40–44.
4. *Харин, Ю.С.* Цепь Маркова s-го порядка с r частичными связями и статистические выводы о ее параметрах / Ю.С. Харин, А. И. Петлицкий // Дискретная математика. – 2007. – Т. 19. – № 2. – С. 109–130.

СТАНДАРТНАЯ ПРОГРАММА РАСЧЕТА ЭТАЛОННЫХ РЕНТГЕНОГРАММ КРИСТАЛЛОВ

***Ф.А. СИТКЕВИЧ*, В.А. ЛИОПО**

Any crystal structure is described in different ways: the model cell, polyhedral model. To calculate the X-ray is necessary to know the coordinates of the atom and atomic scattering amplitude. In this paper we present an algorithm for the calculation of the structure factor on the results of chemical analysis

Ключевые слова: индексы Миллера, структурная амплитуда, рентгенограмма

Рассеяние рентгеновских лучей кристаллом следует начать рассматривать с рассеяния одной его элементарной ячейкой, в которой может находиться несколько атомов разных химических элементов. Каждый атом будет создавать рассеянную волну, амплитуда которой равна атомному фактору рассеяния f_j, в свою очередь зависящему от числа электронов или от порядкового номера данного атома. Очевидно, что сумма волн, рассеиваемых каждым атомом, создаст результирующую волну [1, 2].

Для расчета рентгенограммы необходимо знать координаты атома и его атомную амплитуду рассеяния f_j. Структурная амплитуда $F(h,k,l)$ рассчитывается по формуле:

$$F(h,k,l)=\sum_{j=1}^{N} f_j \exp 2\pi i(\mathrm{h}x_j+\mathrm{k}y_j+\mathrm{l}z_j),$$

где N – число атомов в ячейке, f_j – табличная величина.

Если рассматривать кристалл, у которого часть атомов его идеальной модели замещена другими атомами, то расчет $F(h,k,l)$ существенно усложняется.

В этом случае необходимо выполнить работу по следующему алгоритму:

1.На основе идеальной структурной модели разделить все атомы по различным структурным позициям.

2. Найти число атомов (A_0).

3. Распределить решётки по соответствующим структурным полиэдрам.

4. Из химического состава изучаемого образца выделить число атомов, сумма которых равна А.

5. Найти коэффициент нормировки К из условия: