

УДК 004.896:004.94

И.А. Мошкина, Е.Н. Эгов

ПРОГНОЗИРОВАНИЕ СОСТОЯНИЯ ПРОЕКТА ПО ВРЕМЕННЫМ РЯДАМ МЕТРИК И ВЫЯВЛЕННЫМ АНОМАЛИЯМ¹

Мошкина Ирина Александровна, кандидат технических наук, доцент кафедры «Информационные системы» Ульяновского государственного технического университета. Окончила факультет информационных систем и технологий УлГТУ. Имеет работы в области интеллектуального анализа временных рядов. [e-mail: timina_i@mail.ru].

Эгов Евгений Николаевич, окончил факультет информационных систем и технологий УлГТУ. Ассистент кафедры «Информационные системы» УлГТУ. Имеет статьи в области интеллектуального анализа данных. [e-mail: e.egov@ulstu.ru].

Аннотация

В данной статье рассматривается пример выявления аномалий при анализе временных рядов (ВР) метрик, характеризующих проектную деятельность, для корректировки прогноза состояния проекта. Изложен анализ метрик проектной деятельности. Разработан и реализован алгоритм прогноза на основе нечетких тенденций ВР метрик. Предложен способ выявления аномалий ВР на основе энтропии. Предложена формула вычисления меры энтропии для нечеткого ВР. Алгоритм позволяет учесть зависимость прогнозных значений от мер энтропии. Для прогнозирования используется гипотеза, формируемая на заданный период на основе тенденции. Приведены результаты применения предлагаемого подхода для прогнозирования состояний проекта «FreeNAS9».

Ключевые слова: временной ряд, метрики программного обеспечения, энтропия, прогнозирование.

doi: 10.35752/1991-2927-2019-2-56-67-74

PROJECT STATE FORECASTING ON TIME SERIES OF METRICS AND ON DETECTED ANOMALIES

Irina Aleksandrovna Moshkina, Candidate of Science in Engineering; Associate Professor at the Department of Information Systems of Ulyanovsk State Technical University; graduated from the Faculty of Information Systems and Technologies of Ulyanovsk State Technical University; an author of articles in the field of intellectual analysis of time series. e-mail: timina_i@mail.ru.

Evgenii Nikolaevich Egov, graduated from the Faculty of Information Systems and Technologies of Ulyanovsk State Technical University; Assistant at the Department of Information Systems at Ulyanovsk State Technical University; at author of articles in the field of data mining. e-mail: e.egov@ulstu.ru.

Abstract

The article deals with an example of the anomalies detection when analyzing the time series of metrics that characterize the project activity to adjust the project state forecasting. Project activity metrics are analyzed. A forecasting algorithm based on fuzzy tendency of time series metrics is developed and implemented. Authors suggest a procedure for detecting the time series anomaly based on entropy. A formula for computing the entropy measure for a fuzzy time series is proposed. The algorithm allows to take into account the dependence of the predicted values on the entropy measures. For forecasting, a hypothesis is used, which is formed for a given period on the basis of a trend. The results of the application of the proposed approach for forecasting the state of the project "FreeNAS9" are given.

Key words: time series, software metrics, entropy, forecasting.

¹ Исследование выполнено при финансовой поддержке РФФИ № 18-47-730022, № 18-47-732016, № 18-47-730025, № 19-07-00999.

ВВЕДЕНИЕ

Одной из важнейших задач успешной организации проектной деятельности является эффективное управление массивом проектов за счет прогнозирования их развития. Для того чтобы решить проблему мониторинга состояния проектной деятельности, необходимо иметь инструмент измерения характеристик проектной деятельности, который позволяет отследить динамику совокупности проектов, идентифицировать проблемы и своевременно принять корректирующие регулирующие воздействия. Под управляющим воздействием в первую очередь подразумевается перераспределение ресурсов между проектами [1]. В связи с этим создание средства для управления проектом на основе прогноза значений проектных метрик, представленных в форме временных рядов (ВР), представляется актуальной и важной научно-технической задачей.

Задача анализа и прогнозирования состояний проекта актуальна в связи с накоплением большого количества реализованных продуктов в крупных проектных организациях. Существует довольно много работ, посвященных данной проблеме [2–8].

В статье приведено решение проблемы мониторинга проекта посредством применения средств анализа метрик проекта, извлекаемых из системы контроля версий для построения прогноза и его корректировки на основе мер энтропии.

АНАЛИЗ МЕТРИК ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ

Модель анализа и управления совокупностью проектов в процессе проектной деятельности:

$$\{C_t, R_t, B_t, I_t, F_t, R^{BI}, R^{IF}\}, \quad (1)$$

где C_t – ВР изменений (commits),

R_t – ВР выпусков (release),

B_t – ВР ошибок (bugs),

I_t – ВР улучшений (improvement),

F_t – ВР задач (NewFeature),

R^{BI} – зависимость количества bugs от improvements,

R^{IF} – зависимость новых функциональных свойств от количества усовершенствований (New Features от improvements).

Дискретные ряды представляют данные проекта.

Модель ВР для решения данной задачи построена в два этапа:

1. Дискретный ВР $Y = \{t_i, x_i\}$, $i \in [1, n]$, где t_i – некоторый момент времени, x_i – уровень ВР, преобразован в нечеткий ВР (НВР) $\tilde{Y} = \{t_i, \tilde{x}_i\}$, $i \in [1, n]$, $X = \{x_i\}$, $\tilde{x}_i \in \tilde{X}$, где \tilde{x}_i – нечеткая метка [9].

2. НВР преобразован во ВР нечетких элементарных тенденций. Он обозначен $\tau_i = \text{ETend}(\tilde{x}_i, \tilde{x}_{i+1})$, где ETend – функционал, который реализован на основе операций: TTend – определение типа тенденции и RTend – определение интенсивности тенденции

по специальной лингвистической шкале, построенной для исходного ВР [10].

Несколько четких числовых и нечетких лингвистических ВР получены в результате описанных преобразований (например, ВР параметров нечетких тенденций).

Обобщенная модель НВР нечеткой тенденции представлена [11]:

$$\tau_i = f(\tau_{i-1}, \dots, \tau_{i-d}), \quad (2)$$

где d – фиксированное число, параметр модели;

$\tau_i, \tau_{i-1}, \dots, \tau_{i-d}$ – последовательность нечетких тенденций;

f – некоторая нечеткая зависимость.

Каждая нечеткая тенденция $\tau \in \mathfrak{T}$ НВР представлена структурной моделью [12]:

$$\tau = \langle \tilde{\nu}, \tilde{a}, \Delta t, \mu \rangle, \quad (3)$$

где τ – наименование нечеткой тенденции из множества \mathfrak{T} , $\tau \in \mathfrak{T}$;

$\tilde{\nu}$ – тип нечеткой тенденции, $\tilde{\nu} \in \tilde{V}$, {Падение, Рост, Стабильность};

\tilde{a} – интенсивность нечеткой тенденции, $\tilde{a} \in \tilde{A}$, {Сильное, Среднее, Слабое};

Δt – продолжительность нечеткой тенденции, $\Delta t \in \Delta T$;

μ – функция принадлежности участка НВР, ограниченного интервалом Δt , нечеткой тенденции τ .

Нечеткая тенденция НВР классифицирована в зависимости от Δt :

- элементарная нечеткая тенденция $T \in \mathfrak{T} (\Delta t = 1)$,

- локальная нечеткая тенденция $NT \in N\mathfrak{T}$ ($1 < \Delta t < n - 1$),

- основная (общая) нечеткая тенденция $GT \in G\mathfrak{T}$ ($\Delta t = n - 1$).

ВР нечеткой элементарной тенденции представлен в виде:

$$\begin{aligned} \tilde{\nu}_t &= \text{TTend}(\tilde{x}_t, \tilde{x}_{t+1}), \tilde{a}_t = \text{RTend}(\tilde{x}_t, \tilde{x}_{t+1}), \\ \mu_t &= \min(\mu(\tilde{x}_t), \mu(\tilde{x}_{t+1})), \end{aligned} \quad (4)$$

где $\text{TTend}(\tilde{x}_t, \tilde{x}_{t+1})$ – нечеткое лингвистическое отношение, применяемое для определения сравнительной нечеткой оценки, характеризующей направление изменения (увеличение или уменьшение) между значениями абсолютных нечетких оценок;

$\text{RTend}(\tilde{x}_t, \tilde{x}_{t+1})$ – нечеткое лингвистическое отношение, применяемое для определения сравнительной нечеткой оценки, характеризующей степень различия, «метрическое расстояние» между нечеткими оценками.

Уровни ВР описаны компонентами нечеткой тенденции – тип и интенсивность: $\tau = \{t_i, \tau_i\}$, $i \in [1, n]$, где $\tau = \langle \tilde{\nu}, \tilde{a}, \mu \rangle$.

Метод моделирования нечеткой элементарной тенденции использован для прогнозирования числовых значений и нечетких тенденций состояния проекта

организации в данном продукте. Прогноз строится на основе подтвержденной гипотезы из списка:

Гипотеза 1. Гипотеза сохранения/смены тенденции. Формула прогнозируемого значения при условии сохранения/смены тенденции за предыдущий период рассмотрена:

$$\tau_{t+1} = \tau_t + \tau_p, \quad (5)$$

где τ_{t+1} – прогноз на следующий период времени;

τ_t – реальное значение в момент времени t ;

τ_p – значение тенденции за предыдущий период времени.

Гипотеза 2. Гипотеза устойчивости/неустойчивости тенденции. Формула прогнозируемого значения при условии устойчивости/неустойчивости тенденции, включает расчет скользящего среднего:

$$\tau_{t+1} = \tau_t + G\tau_p, \quad (6)$$

где $G\tau_p$ – значение доминирующей нечеткой тенденции.

Алгоритм формирования прогноза с условием принятия гипотезы 2 истинной:

1. Рассматривается тенденция за предыдущий выбранный период.
2. Выбирается преобладающий кластер тенденции.
3. Вычисляется прогноз по приведенной выше формуле.
4. Формируется тенденция.
5. При одинаковых количествах вхождений тенденций выбирается оптимистический прогноз.
6. Выбирается наибольшее среднее значение тенденций.

Гипотеза 3. Прогнозирование на заданный период на основе нечеткой элементарной тенденции. Этапы алгоритма прогнозирования на период на основе тенденций.

Эксперт задает количество рассматриваемых тенденций за предыдущий период. Например, за полгода – набор тенденций A . Так же эксперт может задать шаблон набора тенденций, для которого известна следующая тенденция:

$$\{\tau_{t_n-m}, \dots, \tau_{t_n-1}, \tau_{t_n}\}. \quad (7)$$

1. Поиск набора A тенденции во всех остальных предыдущих периодах:

$$\left\{ \tau'_{t_{n-l-k}}, \dots, \tau'_{t_{n-l-(k-1)}}, \tau'_{t_{n-l}} \right\}. \quad (8)$$

2. Найден такой набор B , в котором учитывается тенденция C , стоящая после найденного набора B . Формируется прогноз, равный тенденции C :

$$\tau_{t+1} = \tau_t + \tau'_{t_{n-l+1}}. \quad (9)$$

3. Если набора B , который совпадал бы с набором A , не найдено, то поиск набора повторяется. При повторном поиске ищется не полное совпадение с набором A , а с отклонением в одну тенденцию. Это повторяется, пока не будет найден подходящий набор тенденций B [13].

Для выбора наилучшей гипотезы дополнительно вводится энтропийный ВР (ЭВР) [14].

Поиск аномалий ВР на основе энтропии

ЭВР можно представить следующей моделью:

$$\left\{ \tilde{x}_i, H_i^\mu, H_i^T \right\}, \quad (10)$$

где \tilde{x}_i – значение нечеткой метки в i -й точке,

H_i^μ – значение меры энтропии по функции принадлежности в i -й точке,

H_i^T – значение меры энтропии по нечеткой тенденции.

ЭВР позволяет более тонко реагировать на изменения во ВР.

Мера энтропии по функции принадлежности

Мера энтропии по функции принадлежности получается по формуле [15]:

$$H_i^\mu = H(\tilde{A}) + H(\tilde{A}^C), \quad (11)$$

$$\text{где } H(\tilde{A}) = -K \sum_{j=1}^n \mu_{\tilde{A}}(x_j) \ln(\mu_{\tilde{A}}(x_j)), \quad (12)$$

где K – положительная константа,

\tilde{A}^C – дополнение \tilde{A} ,

$\mu_{\tilde{A}}(x_j)$ – значение функции принадлежности точки x к j -му кластеру.

Вычисление функции принадлежности зависит от варианта преобразования ВР в НВР. Например, при фаззификации функция принадлежности может быть задана треугольной функцией принадлежности, имеющей вид:

$$\mu(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b, \\ \frac{c-x}{x-b}, & b < x \leq c, \\ 0, & x \notin \langle a, x \rangle c, \end{cases} \quad (13)$$

где x – значение точки ряда;

$[a, c]$ – диапазон изменения переменной;

b – наиболее возможное значение переменной.

Если НВР формируется с помощью методов кластеризации, то в качестве значений функции принадлежности точек к кластерам берутся аппроксимированные значения из матрицы принадлежности.

Если брать в качестве примера FCM-кластеризацию, то матрица принадлежности высчитывается по следующей формуле:

$$\mu_{j,k} = \frac{1}{\sum_{l=1}^C \left(\frac{\|x_k - c_l\|}{\|x_k - c_j\|} \right)^{\frac{2}{m-1}}}, \quad (14)$$

где $\mu_{j,k}$ – степень принадлежности объекта k кластеру j ,

c_j – вектор центра j -го кластера,

c_j – вектор центра I -го кластера.

При этом если для некоторого кластера $j \|x - c_j\| = 0$, то $\mu_j = 1$, а для всех остальных кластеров равнялась 0.

Мера энтропии по функции принадлежности определяется в 2 шага:

1) Вычисляется значение энтропии по функции принадлежности по формуле (11).

2) На основе полученного значения определяется лингвистическая интерпретация меры энтропии. Значение меры энтропии, близкое к 0, соответствует состоянию «Достоверно». Значение меры энтропии, близкое к максимальному, соответствует состоянию «Неопределенно». В остальных случаях значение меры энтропии соответствует состоянию «Вероятно».

$$\tilde{H}_i^\mu = \begin{cases} \text{Достоверно}, (H_i^\mu \rightarrow 0), \\ \text{Неопределенно}, (H_i^\mu \rightarrow \max), \\ \text{Вероятно}, H_i^\mu \in [0, 1], \end{cases}$$

Полученная на основе функции принадлежности мера энтропии не способна четко фиксировать смену нечетких меток ВР. Мера энтропии показывает, с какой принадлежностью точка будет относиться к метке. При этом, если энтропия близка к максимальному значению, то это говорит о том, что точка находится в «пограничном» положении и может относиться с равной долей вероятности к двум разным нечетким меткам.

Мера энтропии по нечеткой тенденции

Данный метод основывается на предположении, что хаос в системе наступает, когда начинаются отклонения в тенденции показателя. То есть наступает ситуация неопределенности, когда становится сложно предугадать дальнейшее поведение системы. Эту ситуацию можно определить отклонением прогнозного значения тенденции. Оно может быть получено на основе разницы значений тенденции на двух предыдущих точках от фактического значения тенденции в этой точке. Если прогнозные и фактические значения совпадают, то можно говорить о стабильности системы, иначе – о наличии отклонения. Отклонение может быть незначительным или же приводить систему в состояние хаоса и неопределенности [16].

$$H_i^T = \tilde{t}_i^{\text{факт}} - \tilde{t}_i^{\text{прогноз}}, \quad (15)$$

где $\tilde{t}_i^{\text{прогноз}} = \text{GetForecastTrend}(\tilde{t}_{i-1}^{\text{факт}}, \tilde{t}_{i-2}^{\text{факт}})$.

Функция $\text{GetForecastTrend}(\tilde{t}_{i-1}^{\text{факт}}, \tilde{t}_{i-2}^{\text{факт}})$ на основе значений элементарных тенденций в двух предыдущих точках вычисляет динамику тенденции.

$$\Delta\tau_{i-1}^{\text{факт}} = \tau_{i-2}^{\text{факт}} - \tau_{i-1}^{\text{факт}}. \quad (16)$$

Далее определяется код точки $(\tau_{i-1}^{\text{факт}}, \Delta\tau_{i-1}^{\text{факт}})$ по ее положению на координатной плоскости фазовой плоскости.

$$p = \text{CalcCodePoint}(\tau_{i-1}^{\text{факт}}, \Delta\tau_{i-1}^{\text{факт}}). \quad (17)$$

Исходя из имеющихся схем поведения тенденции на фазовой плоскости, выбирается наиболее вероятная и определяется значение элементарной тенденции в следующей точке, после точки $(\tau_{i-1}^{\text{факт}}, \Delta\tau_{i-1}^{\text{факт}})$. Оно и будет прогнозируемым значением элементарной тенденции.

Мера энтропии по нечеткой тенденции определяется по следующему алгоритму:

1. Определяется динамика тенденции в предыдущей точке на основе формулы (16).

2. На основе веса нечеткой тенденции и значения динамики нечеткой тенденции в предыдущей точке вычисляется ее позиция на фазовой плоскости по формуле (17):

$$P_{(i-1)} = \text{CalcCodePoint}(\tau_{i-1}^{\text{факт}}, \Delta\tau_{i-1}^{\text{факт}}).$$

3. Определяется 3 набора точек фазовой плоскости:

а) наиболее вероятные – это точки (как правило, одна точка), которые чаще всего следуют после точки $P_{(i-1)}$: $\omega_{\text{mostexpect}} = \text{Max}(\text{Probability}(P_{i-1}))$;

б) вероятные – это точки, которые тоже следуют после точки P_{i-1} , но они не вошли в первый набор:

$$\omega_{\text{probability}} = \text{Probability}(P_{i-1}) \notin \omega_{\text{mostexpect}}$$

с) аномальные – это все точки, не вошедшие в первые два набора (переход в них не ожидается в нормальных условиях функционирования):

$$\omega_{\text{anomaly}} = \text{AllPoint} \notin \text{Probability}(P_{i-1}).$$

4. Определяется динамика тенденции в текущей точке

$$\Delta\tau_i^{\text{факт}} = \tau_{i-1}^{\text{факт}} - \tau_i^{\text{факт}}.$$

5. Вычисляется точка фазовой плоскости для нечеткой тенденции и динамики в текущей точке ряда:

$$p_i = \text{CalcCodePoint}(\tau_i^{\text{факт}}, \Delta\tau_i^{\text{факт}}).$$

6. Определяется, к какому из трех наборов: $\omega_{\text{mostexpect}}$, $\omega_{\text{probability}}$ или ω_{anomaly} относится полученная точка p_i :

- ЕСЛИ точка относится к набору $\omega_{\text{mostexpect}}$, ТО значение меры энтропии приравнивается к 0, так как полученная точка как раз и ожидалась, т. е. ничего нового мы не узнали;

- ЕСЛИ точка относится к набору $\omega_{\text{probability}}$, ТО значение меры энтропии приравнивается к 0,5, так как полученная точка, хоть и не ожидалась, но и не была чем-то абсолютно новым;

- ЕСЛИ точка относится к набору ω_{anomaly} , ТО значение меры энтропии приравнивается к 1, так как полученную точку не ожидалось увидеть, т. е. в данный момент анализируемая система находится в неизвестном состоянии:

$$H_i^T = \begin{cases} 0, (p_i \in \omega_{\text{mostexpect}}), \\ 0,5, (p_i \in \omega_{\text{probability}}), \\ 1, (p_i \in \omega_{\text{anomaly}}). \end{cases}$$

7. Определяется лингвистическая интерпретация полученного числового значения меры энтропии по нечеткой тенденции H_i^{τ} :

$$\tilde{H}_i^{\tau} = \begin{cases} \text{Стабильность}, (H_i^{\tau} > 0) \cap (H_i^{\tau} < 0,5), \\ \text{Изменение}, (H_i^{\tau} > 0,5) \cap (H_i^{\tau} < 0,7), \\ \text{Аномалия}, (H_i^{\tau} > 0,7). \end{cases}$$

Полученный ЭВР позволяет определить, какую гипотезу лучше использовать. Алгоритм выбора гипотезы, предлагаемый в данной работе, можно охарактеризовать с помощью правил ЕСЛИ-ТО:

1. ЕСЛИ состояние, соответствующее нечеткой метке «Изменение», встречается редко и значение меры энтропии по функции принадлежности в последней точке «Достоверно», ТО принимается гипотеза о сохранении текущей тенденции:

$$\left(\bigcup_{t=1}^T ((H^t == \text{Изменение}) / T < 0,2) \right) \cap (H^T == \text{Достоверно}) \rightarrow \text{Hypothesis 1.}$$

2. ЕСЛИ состояние «Изменение» встречается часто, но отсутствует периодичность, ТО выбирается гипотеза об устойчивости доминирующей тенденции:

$$\left(\bigcup_{t=1}^T (H^t == \text{Изменение}) / T < 0,8 \right) \cap (H^T == \text{Изменение}) \rightarrow \text{Hypothesis 2.}$$

3. ЕСЛИ имеется периодичность (8, 9, 10), ТО выбирается гипотеза 3 (Hypothesis 3) на заданный период.

Если невозможно по мере энтропии и нечеткой тенденции определить гипотезу, то выбор делается на основе значения в последней точке.

4. ЕСЛИ значение меры энтропии по функции принадлежности «Достоверно», а по нечеткой тенденции «Стабильность», ТО выбирается гипотеза о сохранении тенденции:

$$\left(H_{last}^T == \text{Стабильность} \right) \cap \left(H_{last}^{\mu} == \text{Достоверно} \right) \rightarrow \text{Hypothesis 1.}$$

5. ЕСЛИ значение меры энтропии по функции принадлежности «Вероятно» или «Неопределенно» или значение меры энтропии по нечеткой тенденции «Изменение» или «Аномалия», ТО выбирается гипотеза о доминирующей тенденции.

$$\left(\left(H_{last}^T == \text{Изменение} \right) \cup \left(H_{last}^T == \text{Аномалия} \right) \right) \cup \left(\left(H_{last}^{\mu} == \text{Вероятно} \right) \cup \left(H_{last}^{\mu} == \text{Неопределенно} \right) \right) \rightarrow \text{Hypothesis 2.}$$

Полученный данным алгоритмом ЭВР позволяет выявить и скорректировать наиболее верную гипотезу для прогноза.

РЕЗУЛЬТАТЫ ПРОГНОЗИРОВАНИЯ СОСТОЯНИЙ ПРОЕКТА

Для исследований были взяты данные метрик событий открытого проекта «FreeNAS9». Взяты ВР «закрытых» «Bug» и «Feature» [17]. В таблицах 1, 2 представлены результаты анализа и прогнозирования метрик проектной деятельности с исследованием их схожести для выявления зависимости.

В таблице 2 представлен прогноз появления «Bug». Графическое представление прогноза дано на рисунке.

На рисунке представлен вариант прогнозирования с использованием гипотезы, где ось x – значение ВР, ось t – время; «◆» обозначает ВР «Bug»; «▲» обозначает ВР «New Feature»; «●» обозначает прогноз без учета корректировки; «✕» обозначает прогноз после корректировки ВР «Bug» по основной тенденции ВР «Bug»; «■» обозначает прогноз после корректировки ВР «Bug» по основной тенденции ВР «New Feature».

Таблица 1

| Тип метки | Общая тенденция | Доминирующая тенденция | Мера сходства | Корреляция | Интерпретация корреляции |
|-------------|-----------------|------------------------|---------------|------------|--------------------------|
| Bug | Рост | Стабильность | 0,75 | 0,9387 | Сильная |
| New Feature | Рост | Стабильность | | | |

Таблица 2

| ВР зависимый | ВР-предикатор | Гипотеза 1 | Гипотеза 2 | Гипотеза 3 |
|--------------|---------------|-----------------|--------------|-----------------|
| Bug | New Feature | Падение среднее | Рост сильный | Падение среднее |

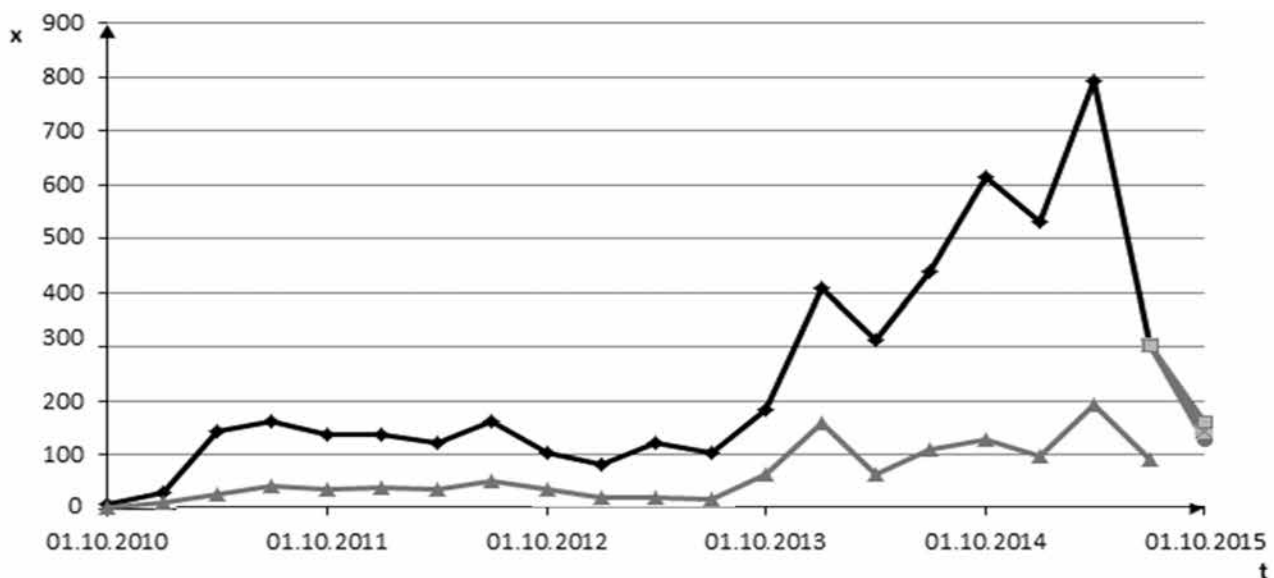


Рис. Прогнозирование появления ошибок в проекте «FreeNAS9» с учетом влияния добавления новых функциональных возможностей в проект (гипотеза прогноза на заданный период)

В таблице 3 представлен ЭВР по метрикам проекта FreeNAS9.

На основе полученных данных можно сделать следующий вывод: гипотеза о сохранении тенденции не подтверждается, так как мера энтропии по нечеткой тенденции в последней точке находится в состоянии «Изменение». Данное состояние, позволяет сделать вывод о смене тенденции во ВР. Наличие периодичности в изменениях статуса проекта указывает на использование гипотезы в заданном периоде.

ЗАКЛЮЧЕНИЕ

Анализ состояния проекта по метрикам, представленным в виде ВР, позволяет формировать прогноз в виде качественной оценки. Меры энтропии по функциональной принадлежности и по нечеткой тенденции являются агрегирующими метриками, которые могут быть использованы при прогнозировании состояний проекта.

Дальнейшими задачами исследования являются расширение числа исходных метрик, характеризующих состояние проекта, исследование применимости предложенного подхода для других типов проектов и сравнение с альтернативными подходами.

Таблица 3

BUG-FreeNAS9 (closed)

| Номер точки | Мера энтропии по функции принадлежности | Мера энтропии по нечеткой тенденции |
|-------------|---|-------------------------------------|
| 1 | Достоверно | Стабильность |
| 2 | Достоверно | Стабильность |
| 3 | Достоверно | Стабильность |
| 4 | Достоверно | Стабильность |
| 5 | Достоверно | Стабильность |
| 6 | Достоверно | Стабильность |
| 7 | Достоверно | Стабильность |
| 8 | Достоверно | Стабильность |
| 9 | Вероятно | Стабильность |
| 10 | Достоверно | Стабильность |
| 11 | Вероятно | Изменение |
| 12 | Достоверно | Стабильность |
| 13 | Вероятно | Стабильность |
| 14 | Вероятно | Изменение |
| 15 | Достоверно | Изменение |

СПИСОК ЛИТЕРАТУРЫ

- Интеллектуальный анализ данных и онтологический подход в управлении проектами / В.С. Мошкин, А.Н. Пирогов, И.А. Тимина, В.В. Шишкин, Н.Г. Ярушкина // Автоматизация процессов управления. – 2016. – № 4 (46). – С. 84–92.
- Гуськов Г.Ю., Наместников А.М., Ярушкина Н.Г. Подход к поиску похожих по структуре проектов, основанный на онтологии языка UML // Радиотехника. – 2017. – № 6. – С. 122–127.
- Гуськов Г.Ю., Наместников А.М. Система управления программными проектами на основе онтологического подхода // Автоматизация процессов управления. – 2016. – № 3 (45). – С. 88–94.
- Гуськов Г.Ю. Алгоритм и программа преобразования проектных диаграмм UML в предметную онтологию OWL // Нечеткие системы, мягкие вычисления и интеллектуальные технологии (НСМВИТ-2017) : тр. VII Всерос. науч.-практич. конф. – СПб. : Политехника-сервис, 2017. – С. 52–57.
- Koukias A., Kiritsis D. Rule-based mechanism to optimize asset management using a technical

documentation ontology // IFAC-PapersOnLine. 2015. Vol. 48 (3). P. 1001–1006.

6. Дородных Н.О. Программная система автоматизации разработки web-сервисов для генерации баз знаний // Системный анализ и информационные технологии (САИТ-2017) : сб. тр. Седьмой Междунар. конф. – М. : Федеральный исследовательский центр «Информатика и управление» РАН, 2017. – С. 222–229.

7. Gavrilova T., Strakhovich E., Leshcheva I. Ontologies for project management teaching // Joint Proceedings of the Work-in-Progress Poster and Invited Young Researcher Symposium at the 18th International Conference on Computers in Education. – Putrajaya, Malaysia, 2010. – pp.7–9.

8. Кроль Т.Я., Харин М.А. Методы решения задачи кластеризации и прогнозирования в электронном архиве // Молодой ученый. – 2011. – Т. 1, № 6. – С. 135–137.

9. Herbst G., Bocklish S.F. Online Recognition of fuzzy time series patterns // International Fuzzy Systems Association World Congress and European Society for Fuzzy Logic and Technology Conference. – Lisbon, Portugal, 2009.

10. Kacprzyk J., Wilbik A. Using Fuzzy Linguistic summaries for the comparison of time series // International Fuzzy Systems Association World Congress and European Society for Fuzzy Logic. – Lisbon, Portugal, 2009.

11. Pedrycz W., Chen S.M. Time Series Analysis, Modeling and Applications: A Computational Intelligence Perspective (e-book Google) // Intelligent Systems Reference Library. – 2013. – Vol. 47. – P. 404.

12. Афанасьева Т.В. Прогнозирование локальных тенденций временных рядов в задачах анализа больших данных // Автоматизация процессов управления. – 2014. – № 4 (38). – С. 42–47.

13. Тимина И.А. Корректировка гипотез прогноза для извлечения знаний о временных рядах // Четырнадцатая национальная конференция по искусственному интеллекту с международным участием (КИИ-2014) : тр. конф., Казань, 24–27 сентября 2014 г. – Казань, 2014. – С. 68–76.

14. Эгов Е.Н., Ярушкина Н.Г., Яшин Д.В. Нечеткое моделирование и генетическая оптимизация временных рядов в интеллектуальной системе технической диагностики // Радиотехника. – 2016. – № 9. – С. 64–71.

15. Ярушкина Н.Г., Воронина В.В., Эгов Е.Н. Применение энтропийной меры в диагностике технических временных рядов // Автоматизация процессов управления. – 2015. – № 2(40). – С. 55–63.

16. Афанасьева Т.В., Заварзин Д.В. Описание подхода к извлечению знаний об аномалиях локальных тенденций временного ряда // Открытые семантические технологии проектирования интеллектуальных систем : матер. VI Междунар. науч.-техн. конф. – Минск, 2016. – С. 305–308.

17. FreeNAS 9 // Bug Tracking System. – URL: <https://bugs.pcbbsd.org/projects/freenas> (дата обращения: 30.10.2017).

REFERENCES

1. Moshkin V.S., A.N. Pirogov, I.A. Timina, V.V. Shishkin, N.G. Yarushkina. Intellectually analysis of ontologicheskii podkhod v upravlenii proektami [Intelligent Analysis of Project and Terminological Metrics in Project Management]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2016, no. 4 (46), pp. 84–92.

2. Guskov G.Iu., Namestnikov A.M., Yarushkina N.G. Podkhod k poisku pokhozhih po strukture proektov, osnovannyi na ontologii iazyka UML [Approach to the Search for Similar Software Project Based on the UML Ontology]. *Radiotekhnika* [Journal Radioengineering], 2017, no. 6, pp. 122–127.

3. Guskov G.Iu., Namestnikov A.M. Sistema upravleniia programmnyimi proektami na osnove ontologicheskogo podkhoda [The Software Project Management System Based on the Ontology Approach]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2016, no. 3 (45), pp. 88–94.

4. Guskov G.Iu. Algoritm i programma preobrazovaniia proektnykh diagramm UML v predmetnuiu ontologiiu OWL [Algorithm and Program for Transformation of UML Project Diagrams into the OWL Domain Ontology]. *Nechetkie sistemy, miagkie vychisleniia i intellektualnye tekhnologii (NSMVIT-2017). Tr. VII Vseros. nauch.-praktich. konf.* [Fuzzy Systems, Soft Computing and Intelligence Technologies (NSMVIT-2017). Proc. of the 7th Russian Sci. Conf. and Workshop]. St. Petersburg, Politehnika-servis Publ., 2017, pp. 52–57.

5. Koukias A., Kiritsis D. Rule-Based Mechanism to Optimize Asset Management Using a Technical Documentation Ontology. *IFAC-PapersOnLine*, 2015, vol. 48 (3), pp. 1001–1006.

6. Dorodnykh N.O. Programmnaia sistema avtomatizatsii razrabotki web-servisov dlia generatsii baz znaniia [Software System of Automation of Web-Service Development for Knowledge Base Generation]. *Sistemnyi analiz i informatsionnye tekhnologii (SAIT-2017). Sb. tr. VII Mezhdunar. konf.* [System Analysis and Information Technologies (SAIT-2017). Proc. of the 7th Int. Conf.]. Moscow, Federal Research Center Informatics and Management, RAS Publ., 2017, pp. 222–229.

7. Gavrilova T., Strakhovich E., Leshcheva I. Ontologies for Project Management Teaching. *Joint Proceedings of the Work-in-Progress Poster and Invited Young Researcher Symposium at the 18th International Conference on Computers in Education*. Putrajaya, Malaysia, 2010, pp.7–9.

8. Krol T.Ia., Kharin M.A. Metody resheniia zadachi klasterizatsii i prognozirovaniia v elektronnom arkhive [Methods for Solving the Problem on Clustering and Prediction in Electronic Archive]. *Molodoi uchenyi* [Young Scientist], 2011, vol. 1, no. 6, pp. 135–137.

9. Herbst G., Bocklish S.F. Online Recognition of Fuzzy Time Series Patterns. *International Fuzzy Systems Association World Congress and European Society for Fuzzy Logic and Technology Conference*, Lisbon, Portugal, 2009.

10. Kacprzyk J., Wilbik A. Using Fuzzy Linguistic Summaries for the Comparison of Time Series. *International Fuzzy Systems Association World Congress and European Society for Fuzzy Logic. Logic and Technology Conference*. Lisbon, Portugal, 2009.
11. Pedrycz W., Chen S.M. Time Series Analysis, Modeling and Applications: A Computational Intelligence Perspective (e-book Google). *Intelligent Systems Reference Library*, 2013, vol. 47, pp. 404.
12. Afanaseva T.V. Prognozirovanie lokalnykh tendentsii vremennykh riadov v zadachakh analiza bolshikh dannykh [Forecasting Time-Series Local Trends in the Big Data Analysis]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2014, no. 4 (38), pp. 42–47.
13. Timina I.A. Korrektirovka gipotez prognoza dlia izvlecheniia znaniia o vremennykh riadakh [Correction of Prediction Hypotheses for Time-Series Knowledge Extraction]. *Chetyrnadtsataia natsionalnaia konferentsiia po iskusstvennomu intellektu s mezhdunarodnym uchastiem (KII-2014). Tr. konf.* [The 14th National Conf. on Artificial Intelligence with International Participation]. Kazan, 2014, pp. 68–76.
14. Egov E.N., Yarushkina N.G., Yashin D.V. Nechetkoe modelirovanie i geneticheskaiia optimizatsiia vremennykh riadov v intellektualnoi sisteme tekhnicheskoi diagnostiki [Fuzzy Modeling and Genetic Optimization of the Time Series in the Intellectual Technical Diagnostics System]. *Radiotekhnika* [Journal Radioengineering], 2016, no 9, pp. 64–71.
15. Yarushkina N.G., Voronina V.V., Egov E.N. Primenenie entropiinoi mery v diagnostike tekhnicheskikh vremennykh riadov [Entropy Application to the Diagnosis of Technical Time Series]. *Avtomatizatsiia protsessov upravleniia* [Automation of Control Processes], 2015, no. 2 (40), pp. 55–63.
16. Afanaseva T.V., Zavarzin D.V. Opisanie podkhoda k izvlecheniiu znaniia ob anomaliiakh lokalnykh tendentsii vremennogo riada [Description of an Approach to the Extraction of Knowledge on Irregular Local Tendency of Time Series]. *Otkrytye semanticheskie tekhnologii proektirovaniia intellektualnykh system. Mater. VI Mezhdunar. nauch.-tekhn. konf.* [Open Semantic Technologies of Intelligent System Design in Proc. of the 6th Int. Sci. and Eng. Conf.]. Minsk, 2016, pp. 305–308.
17. *FreeNAS 9. Bug Tracking System*. Available at: <https://bugs.pcbbsd.org/projects/freenas> (accessed: 30.10.2017).