

**43. ПРОГНОЗИРОВАНИЕ ТРУДОЁМКОСТИ РАЗРАБОТКИ
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ**

Колопенько Е.А., Манкевич Е.Ф.

Аннотация. Работа посвящена актуальной проблеме прогнозирования трудоемкости разработки программного обеспечения. Ресурсы рассматриваются с точки зрения модели СОСОМО Б. Боэма. На основе эмпирических данных зарубежной компании выявляется трудоемкость разработки по среднему уровню модели СОСОМО, проводится регрессионный анализ исходных данных, рассчитывается интервальный прогноз. На основании проведенного анализа был сделан вывод об адекватности модели, даны практические рекомендации использования модели СОСОМО, подтверждена гипотеза о необходимости калибровки исходной модели в новых условиях.

Прогнозирование трудоемкости и продолжительности разработки программного обеспечения является неотъемлемой частью деятельности различных IT-компаний. Ни в одной технологии или управленческой технике не существует универсального метода, увеличивающего на порядок производительность, надежность и простоту. Для большинства методов калибровка в соответствии с областью применения становится обязательным требованием.

Модель конструктивных затрат СОСОМО была описана Барри Боэмом в 1981 году, широко изучена в рамках статистического общества, занимающегося информационными технологиями, и апробирована рядом фирм.

Для проведения регрессионного анализа Боэм использует распределение системы двух случайных величин x_1 (количество строк кода) и x_2 (трудоемкость). При этом величина x_2 представляется как сумма двух случайных величин: $g_2(x_1)$ и поправочного члена ε [1]. Боэм аппроксимирует регрессию степенной функцией. Зависимость трудоемкости от количества строк кода в среднем уровне модели определяется функцией вида:

$$MM = a \times KSLOC^b \times EAF; \quad EAF = \prod_{i=1}^n F_i, \quad (1)$$

где MM – трудоемкость проекта, в человеко-месяцах;

$KSLOC$ – количество строк кода, в тысячах строк;

a, b – коэффициенты уравнения степенной регрессии;

F_i – коэффициент рейтинга i -го фактора стоимости;

n – число драйверов стоимости.

Степенные функции имеют ряд преимуществ, благодаря чему они получили широкое распространение в эконометрических исследованиях: так, коэффициент эластичности для этой зависимости представляет собой постоянную, не зависящую от экзогенной переменной величину. Объективная практическая зависимость трудоемкости от количества инструкций позволяет экономически интерпретировать показатели полученной модели и сделать выбор наилучшего уравнения, исходя из формальных соображений (множественного коэффициента корреляции, средней ошибки аппроксимации, F-критерия Фишера). Регулирующий фактор стоимости интегрирует влияние целого ряда факторов. Кроме того, адаптация модели к разным уровням компетенций персонала и требованиям к аппаратному и программному обеспечению добавляет уверенности в точности модели.

В рамках данного исследования источником проектных данных стала национальная компьютерная консалтинговая и сервисная фирма, специализирующаяся на проектировании и разработке программного обеспечения для обработки данных [2]. Исходя из специфики проектов, коэффициенты a и b равны 3 и 1,12 соответственно [3]. Ввиду ограниченности сведений о проектах и отсутствии достаточных оснований для определения степени влияния факторов стоимости используем номинальные значения факторов ($\forall i F_i = 1$). Получаем следующее уравнение регрессии:

$$MM = 3 \times KSLOC^{1,12}. \quad (2)$$

Построим уравнение регрессии для конкретных эмпирических данных. Получаем модель следующего вида:

$$MM = 2,629 \times KSLOC^{0,815}. \quad (3)$$

Из двух приведенных уравнений регрессии наиболее предпочтительным является уравнение, описываемое формулой 3: наблюдается достаточно тесная связь между случайными величинами (коэффициент корреляции равен 0,785), средняя ошибка аппроксимации, равная 9,215, ниже предельного значения в 15%. Данную модель можно считать адекватной по критерию Фишера, равному 20,947, гипотеза о случайной природе оцениваемых характеристик отклоняется.

Рассчитаем интервальный прогноз для функции регрессии (формула 3). Графическая интерпретация полученных результатов приведена на рисунке 1.

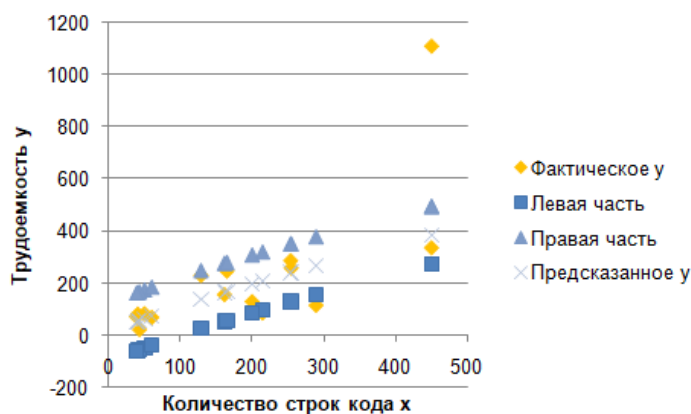


Рисунок 1 – Интервальный прогноз

Использование уравнения регрессии вне пределов обследованного диапазона значений x может привести к значительным погрешностям.

Таким образом, при размере программного продукта в 100 тысяч строк кода для разработки потребуется 112,135 человеко-месяцев, что соответствует сделанному прогнозу. Данная метрика и формула, отражающая выявленные закономерности для взятых эмпирических данных, могут быть использованы для прогнозирования трудоемкости при постоянном уровне человеческих компетенций, сложности и структуре кода в условиях, когда получение всех значений случайных величин генеральной совокупности не представляется возможным. Данное предположение схоже с частной корреляционной зависимостью и связано соответствующими ограничениями.

В настоящее время нет универсального подхода к предварительной адекватной оценке трудоемкости проектов, однако рассматриваемая модель СОСОМО наиболее полно и адекватно учитывает аспекты разработки. При этом все еще существуют некоторые потенциальные ограничения и проблемы. СОСОМО опирается на точные данные и предположения, связанные с программным проектом, соответственно, если эти входные данные неточны или неполны, полученные оценки могут быть ненадежными. Также модель основана на определенных предположениях и параметрах, которые могут неточно отражать уникальные характеристики конкретного программного проекта, что может привести к переоценке или недооценке требуемых ресурсов. Кроме того, настройка различных входных параметров является потенциально трудоемким и утомительным процессом и выдвигает дополнительные требования к компетенциям исполнителей.

Для эффективного использования СОСОМО стоит иметь полное представление о возможностях, ограничениях и допущениях модели. Поскольку СОСОМО опирается на точные данные, связанные с программным проектом, такие как размер, сложность и график, то очень важно собрать эти данные точно, чтобы получить надежные оценки. Для точной оценки следует настроить такие параметры, как уровень опыта команды, зрелость процесса разработки и тип проекта. После получения оценки с помощью СОСОМО рекомендуется проверить результаты на основе исторических данных или контрольных показателей.

Следуя этим рекомендациям, модель СОСОМО можно эффективно использовать для оценки необходимых ресурсов и планирования программного проекта.

Список использованных источников:

1. Корн, Г. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М. : «Наука», 1970. – С. 495.

59-я Научная Конференция Аспирантов, Магистрантов и Студентов БГУИР, Минск, 2023

2. Kemerer, C. F. An Empirical Validation of Software Cost Estimation Models / C. F. Kemerer // Communications of the ACM. – 1987. – Vol. 30, №5. – P. 416-429.
3. Boehm, B. W. Software Engineering Economics / B. W. Boehm. – Prentice Hall, 1981. – 767 p.