

Для побудови класифікатора візьмемо підмножину документів (речення) $\Omega = \{d_1, \dots, d_{|\Omega|}\} \subset D$ та множину категорій SO , при цьому всі елементи Ω вже розподілені за категоріями. Одним із найпростіших підходів для перевірки ефективності роботи класифікатора, є розбиття множини Ω на дві частини $\Omega = Tr \cup Te$ (не обов'язково рівної потужності), причому $Tr \cap Te = \emptyset$. Tr називається тренувальною множиною. Речення з множини Tr будуть використані для побудови (навчання) класифікатора. Te називається тестовою множиною. Після побудови класифікатора елементи множини Te будуть класифіковані. Порівнюючи $\tilde{P}(d_j)$ (результат виданий класифікатором для кожного речення) та істинне значення $\tilde{P}(d_j)$ для кожного $d_j \in Te$, можна оцінити ефективність побудованої моделі класифікатора.

Також використовують більш складний підхід перевірки ефективності, який називається кросвалідація. При такому підході Ω розбивається на k рівних частин $\Omega = Te_1 \cup \dots \cup Te_k$, після чого будується k класифікаторів $\tilde{P}_i (i = 1..k)$. Для побудови \tilde{P}_i в якості тренувальної множини використовується $\Omega \setminus Te_i$, та Te_i використовується для тестування. Результируюча оцінка ефективності класифікатора, є результатом усереднення оцінок ефективностей $\tilde{P}_i (i = 1..k)$.

Список використаних джерел

1. Sebastiani F. A tutorial on automated text categorisation. // Proceedings of the 1st Argentinian Symposium on Artificial Intelligence (ASAI'99). - Buenos Aires, 1999. - p. 7-34.
2. Bing L. Web Data Mining. Exploring Hyperlinks, Contents and Usage Data : / Liu Bing. – Springer, 2006. – 532 с.

УДК 004.652

РАЗРАБОТКА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ВЫПОЛНЕНИЯ ПРОЕКТОВ В ОРГАНИЗАЦИИ

Шовкопляс Ю.В.

Харьковский национальный университет радиоэлектроники

Введение

В настоящее время достаточно большое количество организаций ведет свою работу на основании управления проектами. Основной проблемой, которую приходится решать руководителям проектов, является управление рисками. «Мы живем в эру, в которой принятие рисков вознаграждается, оставляя компании, избегающие рисков, в виде добычи, которая будет разделена между другими» [1].

I. Постановка проблемы

В современной литературе теме управления отдельного проекта уделено много внимания, хотя процесс управления проектами в рамках целой организации недостаточно полно. Таким образом, актуальной является задача формализации процесса моделирования функционирования проектных организаций в условиях риска и неопределенности.

II. Цель работы

Пусть существует некоторая проектная организация, работу которой обобщенно можно представить в виде последовательных фаз проектирования, разработки и тестирования. Проекты поступают в организацию по определенному закону. Процесс выполнения проектов представляет собой спиральную модель, подразумевающую возможность прохождения одним проектом некоторых фаз больше одного раза.

В данной работе предполагается разработка математических моделей проектной организации, выполняющей проекты по спиральной модели, в условиях риска и неопределенности. Необходимо построить математические модели функционирования проектной организации, используя имитационное и аналитическое моделирование. Разработать программное средство реализации полученных моделей и провести сравнительный анализ используемых математических моделей с помощью разработанного программного средства.

III. Разработка математических моделей

Исторически первым сложился аналитический подход к исследованию систем, когда ЭВМ использовалась в качестве вычислителя по аналитическим зависимостям. Анализ характеристик процессов функционирования больших систем с помощью только аналитических методов исследования наталкивается обычно на значительные трудности, приводящие к необходимости существенного упрощения моделей либо на этапе их построения, либо в процессе работы с моделью, что может привести к получению недостоверных результатов. Поэтому в настоящее время наряду с построением аналитических моделей большое внимание уделяется задачам оценки характеристик больших систем на основе имитационных моделей, реализованных на современных ЭВМ с высоким быстродействием и большим объемом оперативной памяти. Основным преимуществом имитационного моделирования по сравнению с аналитическим является возможность решения более сложных задач. В настоящее время имитационное моделирование — наиболее эффективный метод исследования больших систем, а часто и единственный практически доступный метод получения информации о поведении системы, особенно на этапе ее проектирования [2].

Аналитические модели представляется в виде совокупности математических формул и выкладок. Среди достоинств аналитических моделей можно выделить их высокую точность и то, что при достаточно высокой сложности моделируемой системы их использование позволяет существенно снизить аппаратные и временные затраты на моделирование [3].

Для получения основных показателей эффективности функционирования больших систем по ряду причин целесообразно использовать теоретико-вероятностные методы, в частности методы теории массового обслуживания [4].

Если фазы выполнения проектов интерпретировать как обслуживаемые приборы, а проекты — как обслуживаемые заявки, то проектную организацию можно представить в виде системы массового обслуживания. В разрабатываемой системе массового обслуживания все операционные характеристики будут рассчитываться в отдельности для каждой из тех однофазных систем, на которые можно декомпозировать исходную систему, а затем для всей системы будут вычисляться усредненные значения этих характеристик.

Моделирующие алгоритмы могут быть также представлены в виде операторных схем [6]. Имитационную модель проектной организации целесообразно представить в виде операторной схемы.

Выводы

Построены математические модели функционирования проектной организации в условиях риска и неопределенности, используя имитационное и аналитическое моделирование. Проектная организация рассматривалась в виде системы массового обслуживания; разработано программное средство реализации полученных моделей и проведен сравнительный анализ используемых математических моделей с помощью разработанного программного средства. С точки зрения времени проведения моделирования, аналитические методы работают быстрее, чем имитационные. При невысоких частотах возникновения проектов, имитационный метод требует примерно тот же объем машинного времени на выполнение эксперимента, что и аналитический метод, хотя с возрастанием частоты поступления проектов имитационный метод требует гораздо больше времени, чем аналитический. Имитационные методы, в свою очередь, позволяют описать функционирование организации любой сложности и с любыми параметрами.

Список использованных источников

1. ДеМарко Т., Листер Т. Вальсирование с медведями: управление рисками в проектах по разработке программного обеспечения. - Нью-Йорк: «Дорсет Хаус Пабблишинг», - 2003. - 196 с.
2. Советов Б.Я., Яковлев С.А. Моделирование систем. - М.: Высш. шк., - 2001. - 343 с.
3. Основы моделирования сложных систем /Под общ. ред. И.В. Кузьмина - Киев: Вища шк., - 1981. - 360 с.
4. Гнеденко Б.В., Коваленко И.М. Применение теории массового обслуживания к задачам больших систем. — В кн.: Большие системы. Теория, методология моделирования. — М.: Наука, 1971.
5. Бусленко Н. П. Моделирование сложных систем. — М.: Наука, 1988.