

МЕТОДИКА РАСЧЕТА ПРОЕКТНЫХ РИСКОВ НА ОСНОВЕ ПРИМЕНЕНИЯ БАЙЕСОВСКИХ СЕТЕЙ

В статье описана методика расчета рисков на всех этапах жизненного цикла проекта на основе применения Байесовских сетей. Программная реализация данного метода осуществляется с помощью программного пакета Hugin.

The article describes the method of risks calculation on all stages of project life cycle on the basis of application of the Bayes networks. Programmatic realization of this method is carried out by the Hugin software package.

1. Постановка проблемы

Неотъемлемой частью создания качественного программного продукта является составление программного проекта. Анализ и управление проектными рисками становится общепризнанной методологией осуществления проектов и превратилось в неотъемлемую часть ведения бизнеса.

Основной целью управления программными проектами является достижение оптимального качества программного продукта при минимуме затрат, оценка и управление проектными рисками на всех этапах жизненного цикла разработки программного обеспечения.

Можно перечислить несколько преимуществ, которые анализ рисков дает менеджеру: более глубокое понимание специфики проекта, позволяющее создавать более реалистичные планы и бюджет проекта; понимание природы рисков и их потенциальных последствий дает возможность к их распределению между агентами, которые способны лучше ими управлять; возможность оценки резервов, создаваемых для обеспечения рисков, а это, в свою очередь, позволяет уменьшить вероятность срыва проекта [1].

Для успешной реализации перечисленных преимуществ необходимо грамотно уметь анализировать риски на качественном, содержательном уровне с целью их идентификации, выявления причин, оценки потенциальных последствий, разработке стратегии «реагирования» и т.д.

Поскольку сфера разработки программного обеспечения классифицируется как антропоцентрическая система, в которой центральным элементом является человек, то проектные риски неразрывно связаны с условиями его деятельности и влиянием на него факторов внешней и внутренней среды. Следовательно, пути снижения рисков лежат в изучении основных закономерностей влияния факторов среды на интеллек-

туальную деятельность специалистов, задействованных в разработке ПО [2].

2. Анализ исследований и публикаций

Решению данных проблем посвящено значительное количество научных исследований, выполненных в разное время отечественными и зарубежными учеными, среди которых работы В.В. Липаева, Г.Б. Мороза, А.Ф. Кулакова, С.А. Юдицкого, А.М. Вендрова, И. Соммервилла, М. Кантора, Д. Фокса, Д. Лефингвела, Г. Гласа, А. Шаллоуея и др. Стадии подготовки программных проектов, а также рисков, возникающих на данных этапах подробно описали Роберт Т. Фатрелл, Рассел Д. Арчибалд, Г. Дитхелм.

3. Изложение основного материала исследования

По данным статистики, представленной международной компанией Standish Group Chaos, из всех программных проектов, завершенных в 2010 году, только 32% программных проектов являются успешными, 44% являются спорными (имеющими перерасход средств, превышение бюджета, другие недостатки), а 24% являются провальными.

Данная ситуация обусловлена недостаточным вниманием оценке проектных рисков, а также проблемой выбора инструментов планирования.

В настоящее время на рынке представлено значительное количество универсальных программных пакетов для персональных компьютеров, автоматизирующих функции планирования и контроля календарного графика выполнения работ, а также оценки и анализа проектных рисков. Среди наиболее популярных можно привести следующие:

Microsoft Project (Microsoft), OpenProj, TurboProject (IMSI), Spider Project (Технологии управления Спайдер), Project Workbench (Applied Business Technology).

Однако, в современных средствах управления программным проектом задачи численной оценки рисков, обусловленных, например, человеческим фактором, не решаются, а их учет осуществляется на основе интуиции и здравого смысла менеджера [3].

Одним из методов комплексной оценки проектных рисков, является метод вероятностного моделирования с использованием байесовских сетей.

Байесовские сети (БС) представляют собой графовые модели вероятностных и причинно-следственных отношений между переменными в статистическом информационном моделировании. В БС могут органически сочетаться эмпирические частоты появления различных значений переменных, субъективные оценки «ожиданий» и теоретические представления о математических вероятностях тех или иных следствий из априорной информации. Это является важным практическим преимуществом и отличает байесовские сети от других методик оценки риска.

Понятие условной вероятности $P(A|B) = x$ составляет основу байесовского подхода к анализу неопределенности. Совместная вероятности наступления событий A и B дается формулой полной вероятности:

$$P(A, B) = P(A | B) \cdot P(B) \quad (1)$$

Имея в распоряжении информацию о зависимых переменных (следствия), можно определить сравнительные вероятности исходных переменных (причин) при помощи теоремы Байеса.

Пусть имеется условная вероятность $P(A|B)$ наступления некоторого события A при условии, что наступило событие B . Теорема Байеса дает решение для обратной задачи – какова вероятность наступления более раннего события B , если известно, что более позднее событие A наступило.

Пусть A_1, \dots, A_n – полная группа несовместимых взаимоисключающих событий (альтернативных гипотез). Тогда апостериорная вероятность $P(A_j|B)$ каждого из событий A_j при условии, что произошло событие B , выражается через априорную вероятность $P(A_j)$:

$$P(A_j | B) = \frac{P(A_j) \cdot P(B | A_j)}{P(B) = \sum_{i=1}^n P(A_i) \cdot P(B | A_i)} \quad (2)$$

Обратная вероятность $P(B|A_j)$ называется правдоподобием, а знаменатель $P(B)$ в формуле Байеса – свидетельством.

Совместная вероятность является наиболее полным статистическим описанием наблюдаемых данных. Совместное распределение представляется функцией многих переменных в задаче. В общем случае это описание требует задания вероятностей всех допустимых конфигураций значений всех переменных, что мало применимо в реальных задачах. В байесовских сетях, в условиях, когда имеется дополнительная информация о степени зависимости или независимости признаков, эта функция факторизуется на функции меньшего числа переменных:

$$P(A_1, \dots, A_n) = \prod_j P[A_j | Pa(A_j)] \quad (3)$$

где $Pa(A_j)$ – состояния всех переменных-предков для переменной A_j . [4]

Исходя из ранее перечисленных особенностей БС, а так же используя (1), (2) и (3), можно эффективно использовать аппарат БС в задачах разработки модели риска и количественной оценки вероятности его возникновения и материализации, а также влияния на другие риски в проекте.

Основным достоинством использования байесовских сетей в задачах риск-менеджмента является возможность совместного учета количественных и качественных показателей, динамическая обработка поступающей информации, а также явные зависимости между причинами, влияющими на возникновение риска.

Алгоритм использования данного метода включает следующие шаги.

1. Проведение качественного анализа рисков с одновременной оценкой вероятности неблагоприятных событий. При этом необходимо определить степень влияния описываемого риска на проект.

2. Проведение анализа жизненного цикла разрабатываемого ресурса. Необходимо определить, на каких стадиях возникают риски, какие из задач выполняются параллельно, а какие последовательно, чтобы определить влияние одних рисков на другие.

3. Составление правил, которые описывают причинно-следственные связи хода выполнения проекта с учетом рисков.

4. Построение байесовской сети, и соответствующей характеристикам проекта.

5. Задание таблиц условных вероятностей для каждой из нелистьевых вершин байесовской сети.

6. Обучение байесовской сети, проверка адекватности модели.

Для составления БС и таблиц априорных вероятностей для каждого события используются данные, которые могут быть получены от экспертов в области или из опыта многократного выполнения подобных проектов.

Дальнейшее улучшение качества прогнозирования может быть достигнуто путем обучения Байесовской сети на имеющихся экспериментальных данных. Обучение традиционно разделяется на две составляющие – выбор эффективной топологии сети, включая, возможно, добавление новых узлов, соответствующих скрытым переменным, и настройка параметров условных распределений для значений переменных в узлах.

Для расчетов был использован программный пакет Hugin. Hugin является программной реализацией системы принятия решений на основе байесовских сетей доверия. Hugin использует два основных режима работы:

1. Режим редактирования и построения причинно-следственной сети, а также заполнения таблиц условных вероятностей, являющихся количественным описанием БЗ.

2. Режим расчёта вероятностных оценок для принятия решения по всем событиям, входящим в причинно-следственную сеть. Расчёты могут осуществляться как на основе классической теории Байеса, так и на основе методов теории возможностей.

Построим БС, характеризующую поэтапную разработку проекта (рис. 1)

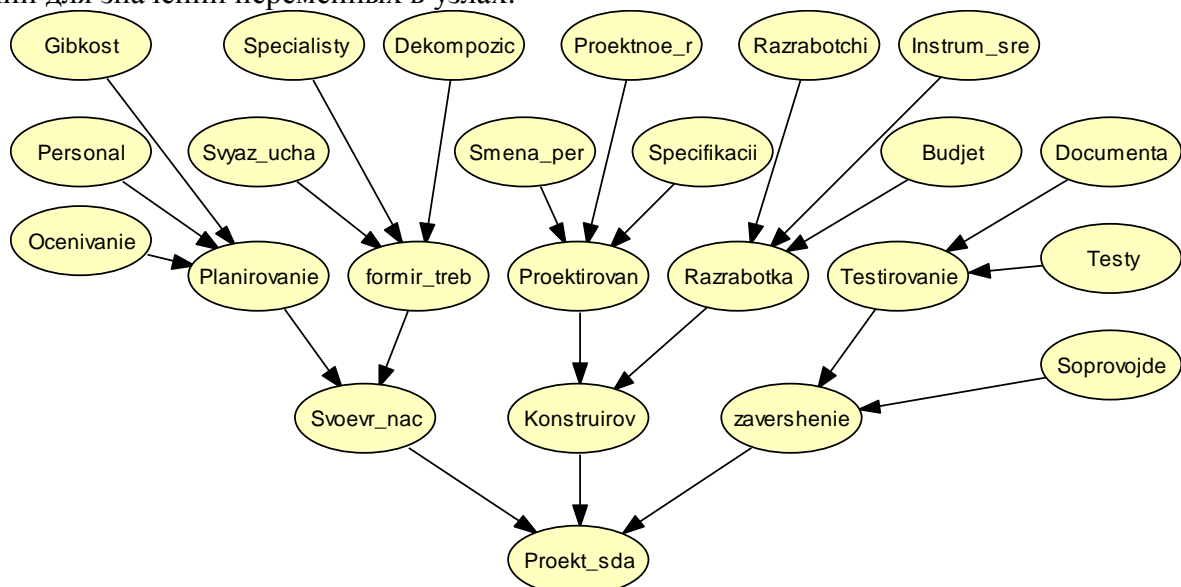


Рис. 1. Байесовская сеть

Зададим таблицы условных вероятностей БС, на всех этапах разработки программного построенных для количественной оценки рисков обеспечения (табл.1):

Табл. 1. Таблица условных вероятностей вершины «Проект сдан»

Вершина	Состояние							
	Успешно		Опоздание		Необходима доработка			
Завершение работ	Успешно				Опоздание			
Конструирование	Успешно		Опоздание		Успешно		Опоздание	
Своевременная подготовка документации	Да	Нет	Да	Нет	Да	Нет	Да	Нет
Проект сдан в срок	0,99	0,85	0,8	0,65	0,85	0,75	0,7	0,01
Проект сдан с опозданием	0,01	0,15	0,2	0,35	0,15	0,25	0,3	0,99

Для того, чтобы дать вероятностную оценку риску необходимо использовать функцию распределения вероятностей. В окне состояний отобразится, что вероятность возникновения события «Проект сдан с опозданием» с учетом условных вероятностей причин равна 10,29%. Можно принять это как максимальную допустимую границу риска и сравнивать полученные результаты с ней. Если они не будут превышать

максимальную допустимую границу, рекомендуется их принимать, иначе рекомендуется исследовать модель риска и внести в проект резервы времени на преодоление последствий материализации риска.

Для нахождения нижней границы значения искомого риска используется распределение аксимумов. Наименьшая вероятность

возникновения риска «Проект сдан с опозданием» составляет 2,13%.

На основе анализа факта срыва сроков проекта, можно сделать предположение, что с вероятностью 47% это произойдет из-за задержки на начальном этапе, с вероятностью 42% – на этапе конструирования, и с вероятностью 48%, что проблемы возникнут при тестировании проекта.

Для того, чтобы найти наиболее вероятную комбинацию состояний всех вершин, необходимо вновь использовать распространение максимумов. После пересчета получим новое распространение в окнах отображения сети и вершин. При этом каждое из состояний вершин, имеющее значение 100%, будет принадлежать к наиболее вероятной комбинации состояний. Для рассматриваемого примера будет получена одна уникальная комбинация, при которой наиболее вероятным является, что проект будет сорван тогда, когда событие «Завершение работ» находится в состоянии «Необходима доработка».

Рассчитаем вероятность наиболее вероятной комбинации состояний, полученных при наличии факта, что проект будет сдан с опозданием. Эта вероятность записывается в виде:

P (Своевременная подготовка документации = «успешно», Конструирование завершено = «успешно», Завершение работ = «необходима доработка» | Проект сдан = «с опозданием») и для ее определения можно воспользоваться формулой:

$$P(A, B) = P(A | B) P(B), \quad (4)$$

переписав ее в виде:

$$P(A | B) = P(A, B) / P(B), \quad (5)$$

Если выбрать состояние «с опозданием» для вершины Проект сдан и сделать расчет распространения сумм, то вероятность пребывания вершины в указанном состоянии будет равно 10,29%. Определив чему равно P (Проект сдан = «с опозданием»), необходимо определить чему равно P (Своевременная подготовка документации = «успешно», Конструирование завершено = «успешно»),

Завершение работ = «необходима доработка», Проект сдан = «с опозданием»). Это значение равно 4,9%. Теперь можно рассчитать искомую вероятность, которая будет равна:

$$P(A, B) / P(B) = 0.049 / 0.103 = 0.476 \quad (6)$$

Таким образом, на основе поступившего факта срыва сроков проекта экспертная система делает вывод о том, что наиболее вероятной является ситуация, связанная с тем, что на завершающем этапе проект потребует доработки. При этом вероятность такой ситуации равна 0.476.

4. Выводы

В статье предложена модель оценки рисков на всех этапах создания программного обеспечения. Байесовские сети предоставляют удобный аппарат для исследования рисков, составления их модели и количественной оценки. Модель риска удобно представлять в виде ориентированного графа, это позволяет как можно точнее исследовать составляющие риска и включить максимальное количество факторов, влияющих на вероятность возникновения риска.

В байесовских сетях доверия есть возможность использовать как вероятности, полученные опытным путем, или из опыта многократного использования системы, так и экспертные оценки, что позволяет использовать в процессе составления модели риска знания эксперта-специалиста, выраженные в виде предположений.

Модель, построенная в виде байесовской сети доверия – динамична, она легко подстраивается под полученные свидетельства. Это позволяет вносить в систему данные, полученные на каждом этапе выполнения проекта. Усовершенствованные таким образом модели рисков могут использоваться в других аналогичных проектах.

Список литературы

1. Фатрелл Р. Управление программными проектами / Р. Фатрелл, Д. Шафер, Л. Шафер; пер. с англ. А. Бойко, А. И. Мороза, А. П. Сергеева. – М. : Издательский дом «Вильямс», 2003. – 1136 с.
2. Руководство по своду знаний по управлению проектами PMBOK / Project Management Institute, Inc., 2009. – 388 с.
3. Арчибальд Р. Управление высокотехнологичными программами и проектами / Р. Арчибальд; пер. с англ. Ю. П. Рогач, М. М. Сердюк. – М: ДМК-Пресс, 2002. – 460 с.
4. Тулупьев А. Л. Байесовские сети: Логико-вероятностный подход / А. Л. Тулупьев, С. И. Николенко, А. В. Сироткин – СПб. : Наука, 2006. — 607 с.