

ДРЕВОВИДНЫЕ КАРТЫ ДЛЯ ПОВЫШЕНИЯ КАЧЕСТВА ПОДДЕРЖКИ РЕШЕНИЙ

М. Ю. Михеев¹, О. В. Прокофьев², И. Ю. Семочкина³

^{1, 2, 3} Пензенский государственный технологический университет, Пенза, Россия
¹ mix1959@gmail.com, ² prokof_ow@mail.ru, ³ ius1961@gmail.com

Аннотация. *Актуальность и цели.* В процессе эволюции компьютерных систем поддержки принятия решений (*Decision Support System, DSS*) «классические» деревья решений были дополнены новыми способами отображения данных – древовидными картами. Используя иерархическую структуру, древовидные карты обеспечивают осмысленно организованные отображения информации большого объема, что дает преимущества при выработке решения по большим данным. Собирая и обобщая опыт применения обычных деревьев и древовидных карт, авторы поставили цель выработать рекомендации по применению средств визуализации данных в прикладных областях, повышающему качество поддержки решений. *Материалы и методы.* Если обычный подход нашел эффективное применение в машинном обучении, то древовидные карты ориентированы на применение «вручную» лицами, принимающими решения. В отличие от первого подхода, опирающегося на объективные числовые критерии оценки качества, эффект от применения древовидных карт субъективен и менее очевиден. Для второго подхода применены данные опросов пользователей во время решения профессиональных задач, связанных с анализом web-страниц «вручную» в процессе полного факторного эксперимента. *Результаты и выводы.* Среднее время решения задач и экспертные оценки позволили сделать выводы о преимуществах древовидного представления данных и потенциальной возможности повышения качества поддержки принимаемых решений. В заключении обоснован вывод о необходимости комплексного применения средств визуализации данных и ряда других технологий в продуктах второго поколения DSS 2.

Ключевые слова: деревья решений, древовидные карты, качество визуализации, качество решения, экспериментальные данные

Для цитирования: Михеев М. Ю., Прокофьев О. В., Семочкина И. Ю. Древовидные карты для повышения качества поддержки решений // Надежность и качество сложных систем. 2021. № 1. С. 76–86. doi:10.21685/2307-4205-2021-1-8

TREEMAPS TO IMPROVE QUALITY OF SUPPORT OF DECISIONS

M.Yu. Mikheev¹, O.V. Prokofev², I.Yu. Semochkina³

^{1, 2, 3} Penza State Technological University, Penza, Russia
¹ e-mail: mix1959@gmail.com, ² prokof_ow@mail.ru, ³ ius1961@gmail.com

Abstract. *Background.* In the process of evolution of Decision Support System (DSS) computer systems, the "classic" decision trees were supplemented with new ways of displaying data – tree maps. Using a hierarchical structure, tree maps provide meaningfully organized displays of large amounts of information, which gives advantages in developing solutions for big data. *Materials and methods.* Gathering and summarizing the experience of using ordinary trees and tree maps, the authors set a goal to develop recommendations on the use of data visualization tools in applied areas, which improves the quality of decision support. If the usual approach has found effective application in machine learning, then tree maps are oriented to manual use by decision makers. In contrast to the first approach, which relies on objective numerical criteria for assessing quality, the effect of using tree maps is subjective and less obvious. For the second approach, user survey data were applied during the solution of professional tasks related to the analysis of web pages "manually" in the process of a full factorial experiment. *Results and conclusions.* The average time for solving problems and expert estimates allowed us to draw conclusions about the advantages of a tree-like presentation of data and the potential for improving the quality of support for decisions made. In conclusion, the article substantiates the conclusion about the need for the integrated use of data visualization tools and a number of other technologies in the second generation DSS 2 products.

Keywords: decision trees, treemaps, quality of visualization, quality of solution, experimental data

For citation: Mikheev M.Yu., Prokofev O.V., Semochkina I.Yu. Treemaps to improve quality of support of decisions. *Nadezhnost' i kachestvo slozhnykh sistem = Reliability and quality of complex systems*. 2021;1:76–86. (In Russ.). doi:10.21685/2307-4205-2021-1-8

Введение

Дерево решений используется в системах поддержки принятия решений как карта возможных результатов ряда связанных вариантов. Это позволяет отдельному лицу или организации сопоставлять возможные действия друг с другом на основе их затрат, вероятностей и выгод. Применение деревьев решений в традиционном смысле неоднократно рассматривалось в литературе по построению DSS [1]. Их можно использовать либо для неформального обсуждения, либо для составления алгоритма, который математически предсказывает лучший выбор [2–9]. Дерево решений обычно начинается с одного узла, который разветвляется на возможные результаты. Каждый из этих результатов приводит к дополнительным узлам, которые переходят в другие возможности, что придает ему древовидную форму. Имея полное дерево решений, можно начать анализ решения в ситуации, с которой столкнулись. Примером современной реализации «классического» подхода являются диаграммы, построенные в среде разработки Lucidchart [10].

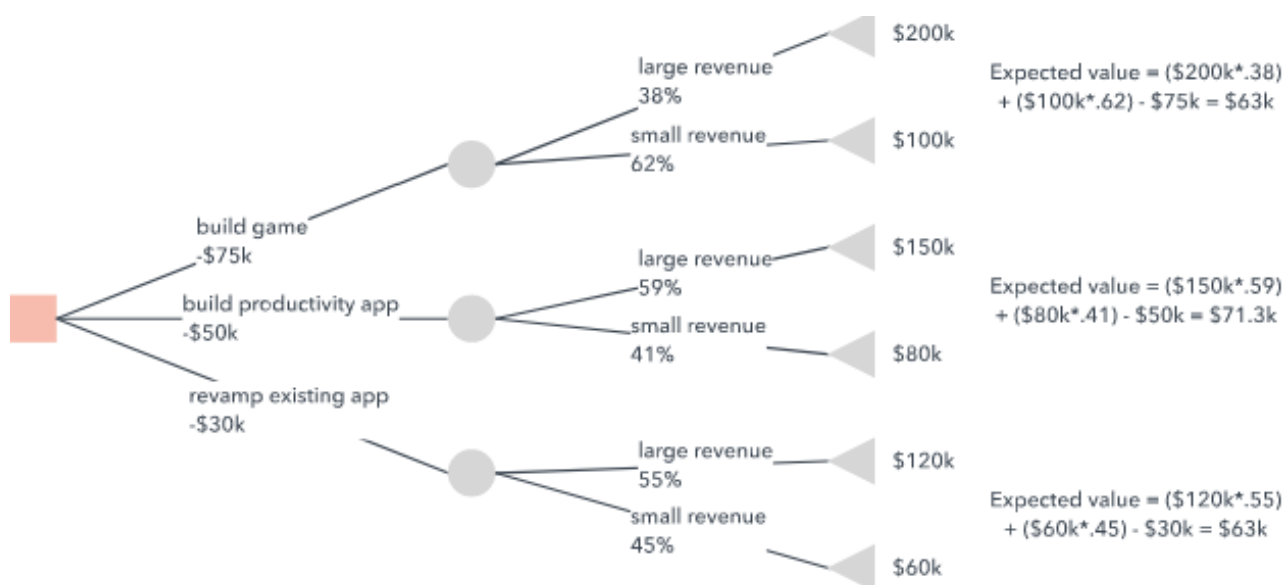


Рис. 1. Пример дерева решений в среде Lucidchart

При определении того, какой результат является наиболее желательным, важно учитывать предпочтения лица, принимающего решения. Например, некоторые могут предпочесть варианты с низким уровнем риска, в то время как другие готовы пойти на риск для большей выгоды. Когда нужно использовать дерево решений с сопутствующей вероятностной моделью, можно применить ее для расчета условной вероятности события или вероятности того, что это произойдет, если произойдет другое событие. Для этого начинают с начального события, затем нужно следовать по пути от этого события к целевому событию, умножая вероятности каждого из этих событий. Таким образом, дерево решений может использоваться как традиционная древовидная диаграмма, которая отображает вероятности определенных событий. Расчет и визуализация в DSS решений с помощью классического вероятностного подхода были только началом для нового этапа, на котором для выработки решений стали применяться неявные данные, полученные в процессе добычи данных, а также данные, выявленные в больших объемах менее информативных данных. Целью работы являлись выявление, систематизация древовидных карт, появившихся в новом поколении продуктов DSS 2 [11], оценка карт в отношении качества визуализации данных для лиц, принимающих решения (ЛПР), выработка рекомендаций по областям применения, ставшими доступными на основе изложенного выше «классического» варианта.

Постановка задачи

Современной разновидностью диаграмм для ЛПР являются древовидные карты, использующие метод визуального представления данных, при котором площадь и цвет прямоугольников представляют две разные метрики. Древовидная карта также может использовать вложенные прямоугольники для представления иерархических отношений. Графическое представление вложенных групп может влиять на понимание пользователями древовидной карты. Существует два распространенных способа представления иерархии в древовидных картах: «простой» подход (рис. 2), который использует плавающий заголовок в центре каждой группы, и подход «заголовки» (рис. 3), использующий отдельный заголовок в прямоугольнике для каждой подгруппы. В обоих случаях при наведении курсора мыши на заголовок группы выделяется граница группы и появляется всплывающая подсказка с общей статистикой группы. Несколько встроенных уровней иерархии приведут к появлению нескольких плавающих заголовков или строк заголовка, что влияет как на видимость, так и на выбор данных. Плавающие заголовки могут скрывать значительную часть нижележащих прямоугольников и другие плавающие заголовки, в то время как подход заголовка использует пространство, которое иначе могло бы показать данные. Количество прямоугольников, представленных в древовидной карте, может влиять на способность пользователей идентифицировать и понять закономерности в данных. Здесь изображены два набора данных различного размера: на рис. 2 показана древовидная карта с 100 ячейками, а на рис. 3 показана одна с 1000 ячейками [12].



Рис. 2. Карта дерева «простой» иерархии с небольшим набором данных (Treemap-Classic)

По сравнению с таблицами древовидные карты должны обеспечивать лучший общий обзор и более тесную интеграцию между измерениями данных. Это преимущество ожидается в задачах, которые требуют некоторой степени грубого сравнения между прямоугольниками и (или) группами. Когда требуется конкретная идентификация, особенно в небольших таблицах, разумно ожидать одинаковой производительности между древовидной картой и табличными представлениями одних и тех же данных. Большие наборы данных потребуют отнимающей много времени прокрутки в таблицах, но, возможно, могут вызвать большее «затенение» данных в древовидных картах. В общем случае, оценка полезности и удобства древовидных карт касается положения узлов в иерархиях, отношений между узлами и группами и идентификации конкретной информации (на основе области,

цвета или того и другого). Дополнительные факторы, которые могут повлиять на сложность задач для древовидной карты, включают: сложность идентификации и сравнения; количество уровней иерархии, пройденных в задачах сравнения; потребность использовать площадь, цвет или обе метрики для решения задачи. Из этого следует, что оценка качества визуализации данных и принимаемых по ним решений требует применения методов формирования группы экспертов, построения тестовых наборов данных, сбора и обработки экспертных оценок.

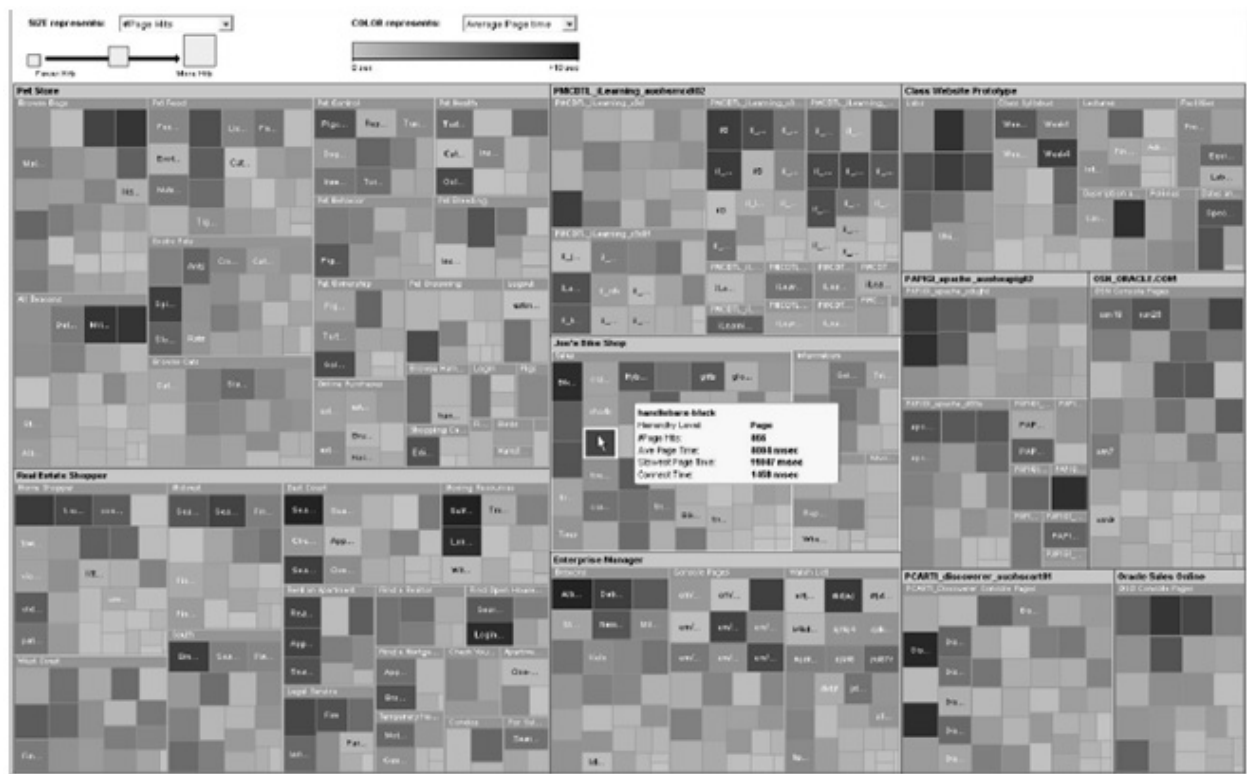


Рис. 3. Карта дерева иерархии «заголовки» с большим набором данных (Treemap-Titlebar)

Эксперименты и обработка результатов

Рассмотрим тестовые данные и организацию эксперимента. Целью экспериментов [12] являлось исследование эффективности и субъективных впечатлений пользователей-ЛПР, использующих древовидные карты и таблицы при идентификации, сравнении и анализе иерархически организованных данных. К управляемым факторам относились древовидная карта и представление данных таблицы, размер набора данных и иерархическое представление древовидной карты (Dataview). Дополнительные факторы включали эффекты обучаемости и эффекты сложности задач. Привлечено 10 экспертов, выдано восемь заданий, в каждом из них имелось шесть условий. Субъективные рейтинговые шкалы были построены после каждого набора из восьми заданий (Task Category), а также проводились собеседования для оценки субъективных впечатлений после выполнения всех заданий. Полный факторный эксперимент использовал факторное пространство 3×2 , соответственно, три способа отображения данных (Treemap-Classic на рис. 2, Treemap-Titlebar на рис. 3) и два различных по размеру набора данных (Large, Small). Восемь задач были заданы для каждого из приведенных выше шести условий (подробности в табл. 1). Они варьировались по сложности от базовых задач идентификации до задач сравнения и анализа. Восемь заданий были представлены в том же порядке в каждом из шести условий, но с разными вопросами. Шесть условий (определяемых представлением данных и размером набора данных) были выровнены по трудоемкости для проведения оценки изучаемости с помощью трех ограничений: каждое условие появлялось примерно одинаковое количество раз, таблицы чередовались с древовидными картами в порядок просмотра и большие наборы данных перемежались с небольшими наборами данных в порядке просмотра.

Таблица 1

Категории задач и примеры

Задачи	Тип задачи	Типичные примеры
1, 2, 5	Идентификация или подсчет (Ident)	Каково было среднее время на странице при аренде квартиры?
3, 4	Сравнение, используя один или несколько критериев (Compare)	Какие страницы имели самое медленное среднее время страницы и наименьшее количество хитов?
7	Расширенное сравнение (Compare2)	Какая транзакция содержала больше всего страниц, которые имеют более 900 просмотров страниц?
6, 8	Открытый анализ (Analysis)	Основываясь только на данных перед вами, каковы (если таковые определяются) отношения между временем страницы-хита и средним временем просмотра страницы?

Три вида данных включали в себя древовидную карту с классическим представлением иерархии (Treemap-Classic или ТМ-С, рис. 2), древовидную карту с представлением иерархии заголовков (Treemap-Titlebar или ТМ-Т, рис. 3) и несортированную таблицу с иерархией, представленной отступом и цветом заголовка (TABL). Таблицы были отсортированы, чтобы обеспечить базовую линию для сравнения производительности на древовидных картах. Каждый набор данных содержал три уровня иерархии: веб-приложения (например, «Магазин велосипедов Джо»), транзакции (например, «Продажи») и страницы (например, «Педали»). Как древовидные карты, так и таблицы позволили участникам свободно выбирать из списка две метрики производительности из списка: количество просмотров страницы, среднее время страницы, самое медленное время страницы и время соединения. Они были представлены в двух выпадающих списках как в таблицах, так и в древовидных картах. Результаты эксперимента, выражающие зависимость среднего времени работы над задачей от значений факторов эксперимента, представлены на рис. 4,а–в, рейтинговые оценки способов визуализации данных по критериям качества показаны на рис. 4,г.

Дисперсионный анализ времени завершения задачи включал условия для каждого способа представления данных (ТМ-Т, ТМ-С, Tabl) номера задачи, порядка просмотра (Viewing Order) и размер набора данных (Dataset). Визуальное представление данных сильно повлияло на время выполнения задачи: задачи ТМ-Т выполнялись на 30 % быстрее (78 с), чем представления ТМ-С (108 с) или TABL (111 с) (критерий Фишера $F = 12$, уровень значимости $p < 0,001$). Увеличение времени завершения от малого к большому набору данных для представления TABL было значительно длиннее аналогичного древовидного (рис. 4,а).

Существенные различия во времени выполнения были обнаружены между восемью задачами в условии ($F = 16,5$, $p < 0,001$). Как показано на рис. 4,б, формат ТМ-Т имел явное преимущество во всех этих категориях задач, сгруппированных в три категории: идентификация значений или подсчет листовых узлов (задачи 1, 2, 5); сравнение по всем наборам данных и более открытый анализ, объединяющий два критерия (задачи 3, 4, 6, 8); специальная задача сравнения, требующая сложного сравнения групп на основе критериев более низкого уровня (задача 7).

Схема уравнивания позволила оценить доступность древовидных карт для изучения. Порядок просмотра значительно повлиял на время выполнения задачи (рис. 4,в, где $F = 8,3$, $p < 0,001$), при этом первое ознакомление с данными участников (153 с) значительно медленнее, чем второе (121 с), которое было значительно медленнее, чем остальные представления (73–86 с). В трех представлениях данных асимптотическая производительность обычно достигалась на третьем или четвертом наборе данных. В целом восприятие таблицы и двух древовидных карт было приблизительно одинаковым, хотя у нескольких экспертов возникли некоторые трудности с интерпретацией ТМ-С в первой позиции просмотра.

Всего было выполнено 14 подсказок и 40 ошибок в 480 завершенных испытаниях (10 участников $\times$ 6 просмотров $\times$ 8 задач). Помощь предоставлялась либо для предотвращения выполнения участником действия, которое препятствовало бы завершению задачи, либо для исправления неверного восприятия участником, которое препятствовало бы завершению задачи. Подсказки были равномерно распределены по представлениям данных и задачам, но 86 % подсказок потребовались,

когда участники использовали большой набор данных. Ошибки были в равной степени распределены по трем способам представления данных, причем большинство возникало в задачах 7 и 3 (57 и 25 % ошибок соответственно). Две трети ошибок в этих двух задачах возникли при использовании большого набора данных. Это были также две самые медленные задачи, подтверждающие сложность проведения сравнений между уровнями иерархии.

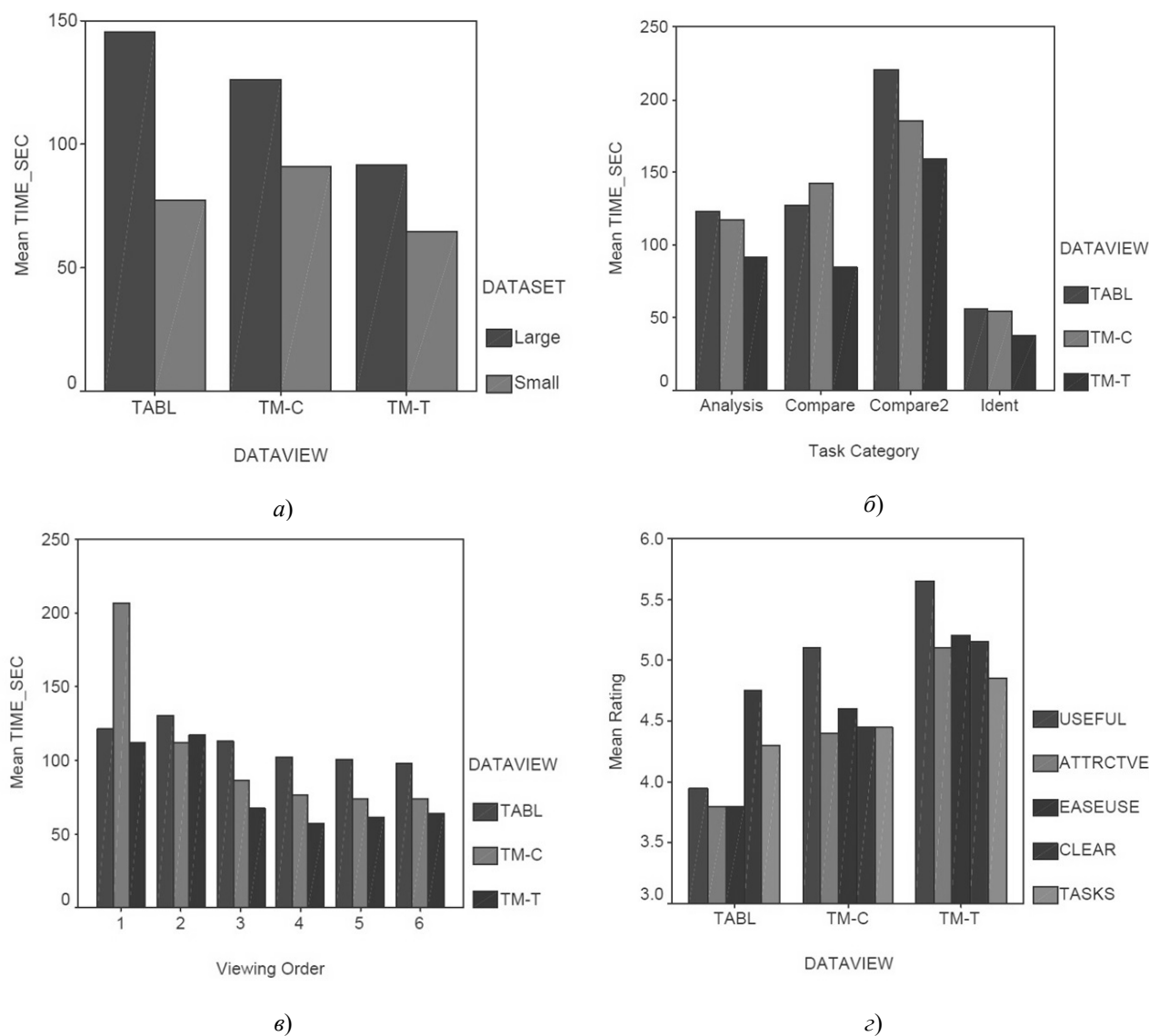


Рис. 4. Результаты экспериментов:

- а – время выполнения задачи, представление данных и размер набора данных;
- б – время выполнения задач по категориям задач и видам представления данных;
- в – время выполнения задачи в порядке просмотра и представления данных;
- г – средний рейтинг пользователей по данным в диапазоне 1–7

Выполнив восемь заданий в каждом условии, участники заполнили пять 7-балльных рейтинговых шкал, касающихся их опыта с этим условием. В каждом случае оценка 1 была отрицательной (например, «совсем не полезно» или «очень сложно»), тогда как оценка 7 была положительной (например, «очень полезна» или «очень проста»). Средние оценки по конкретным вопросам и видам представления данных показаны на рис. 4,г. Вопросы «Насколько это было привлекательно?» и «Насколько легко это было использовать?» дали очень похожие результаты, с общими оценками для TM-T (в среднем 5,2) выше, чем для TM-C (4,6), что, в свою очередь, было выше, чем представление TABL (3,8). При оценке «Насколько это было полезно?» респонденты считали, что две древовидные карты (5,1–5,6) намного превосходят таблицу (3,9). Вопрос «Насколько ясным и понятным это

было?» произвел аналогичные оценки для трех представлений данных, с 4,5–5,0. На вопрос «Как бы вы оценили вашу способность выполнять сегодняшние задачи?» TABL и TM-C были очень похожи (4,4) и были умеренно ниже, чем для TM-T (4,8).

Оценка качества поддержки решений

Это исследование показало, что пользователи (системные администраторы и сетевые администраторы, привлеченные в качестве экспертов) быстро изучают древовидные карты, которые потенциально превосходят табличные представления для задач идентификации, сравнения и анализа. Формат иерархии заголовков TM-T для древовидных карт был предпочтительнее формата TM-C, который может скрыть основную информацию. Преимущество во времени и в субъективных предпочтениях для древовидных карт заметно как для небольших, так и для больших наборов данных.

Большинство решений для визуализации информации, включая древовидные карты, требуют времени для ознакомления и изучения, прежде чем пользователи смогут эффективно использовать их для решения проблем. Неподготовленные участники проходили обучение в первых опытах, независимо от того, обеспечивали ли представление данных древовидную карту или таблицу. Наиболее быстрое изучение было выявлено по среднему времени завершения в случаях древовидной карты. Следовательно, древовидные карты могут быть быстро изучены теми, кто совершенно не знаком с этой концепцией, и это обучение происходит довольно быстро. Включение четких условных обозначений для размера и цвета, а также обширная информация о подсказках для перемещаемых элементов добавили возможность изучения древовидной карты.

Четыре категории задач были включены в экспериментальные задачи этого исследования. Задачи идентификации были выполнены наиболее быстро при всех трех представлениях данных. Задачи сравнения и открытого анализа выполнялись медленнее, но по времени были похожи друг на друга. Задачи расширенного сравнения на уровне страниц были выполнены медленнее всех из трех представлений данных. Представление TM-T превосходило по времени завершения оба других представления в каждой из этих категорий, в то время как TM-C и TABL обеспечивали аналогичную производительность пользователя в целом. Более высокая производительность TM-T также была отражена в рейтингах пользователей, которым было проще пользоваться, они более привлекательны и понятнее, чем другие представления данных.

Производительность по задачам идентификации была намного более похожа между тремя представлениями данных, чем у более сложных задач сравнения. В задачах идентификации участники либо просматривали строки и заголовки в таблице, либо читали информацию всплывающей подсказки в древовидных картах. Основные различия в производительности задач здесь заключались в том, что: древовидные карты содержали данные, видимые без какой-либо прокрутки; участникам часто приходилось прокручивать верхнюю часть таблицы, чтобы определить свои параметры метрики столбца. Представления древовидной карты позволили проводить более точные сравнения между страницами и группами.

Однако ошибки все еще были очевидны, потому что: некоторые пользователи древовидной карты оценивали размер групп по количеству содержащихся прямоугольников, а не по метрике, которая была отображена в области прямоугольника. Таким образом, области группы (веб-приложения и транзакции) отражают количество прямоугольников, а не выбранную метрику. Другая ошибка: пользователи не всегда искали данные до самых краев древовидной карты, пропуская некоторые важные данные.

Различная плотность данных между большими и маленькими по количеству данных древовидными картами влияла на полезное поле зрения участников (Useful field of view, UFOV), то есть на круговую область, в которой указано, из какого источника извлекается информация. UFOV может быть оценен по расстоянию между двумя последовательными фиксациями глаза в задаче поиска; эти расстояния фиксации уменьшаются с увеличением визуальной плотности фоновой информации. Следовательно, более плотные древовидные карты, вероятно, вызывают меньшее UFOV и требуют большего количества фиксаций, что, в свою очередь, может привести к большему количеству пропущенной информации. Представление TM-T превосходило как по объективной производительности, так и по субъективным предпочтениям представление TM-C. Это произошло, несмотря на то, что для формата TM-T была зарезервирована область экрана для имен групп в заголовках, оставляя меньше места для отображения данных конечного узла. В представлении TM-C участники

не всегда замечали, что шрифты для заголовков транзакций были меньше, чем для заголовка веб-приложения, что могло привести к увеличению количества ошибок и времени выполнения.

Потенциальная возможность дальнейшего развития древовидных карт выявлена в ходе высказывания предложений пользователей, участвовавших в эксперименте. Многие участники предложили дополнительные функции древовидной карты, чтобы помочь в их задачах мониторинга сети. Определяемый пользователем порог, после которого конечные узлы станут красными, позволит им более легко отслеживать проблемные сетевые объекты. Автоматическое обновление, генерируемое системой, требуется через определенные администратором интервалы, например, каждые 30 мин. Фильтрация по одной или нескольким конкретным метрикам позволит быстро анализировать проблемы. Некоторым также требуется предоставление исторического представления метрики данных, чтобы лучше диагностировать проблемы. Некоторые участники также отметили, что непрерывная цветовая шкала не нужна, поскольку дискретных значений цвета достаточно для большинства задач мониторинга и диагностики.

Древовидные карты превосходили несортированные представления таблиц в задачах идентификации, сравнения и анализа в этом исследовании. Это преимущество возросло, когда две метрики информации были объединены, и это дает возможность обнаруживать тенденции. Здесь представление иерархии заголовков превосходило классическое представление с плавающим заголовком из-за меньшего количества затуманивания и путаницы, но особые проблемы дизайна, такие как шрифты и цвета, имеют решающее значение при визуализации данных.

Обсуждение и рекомендации

Эволюция методов визуализации иерархических данных привела к созданию новых способов их отображения, которые могут одновременно применяться с известными способами, взаимно дополняя друг друга.

Рассмотрим преимущества и недостатки «обычных» деревьев решений в аналитике. Деревья решений остаются популярными, так как их легко воспринимать, а также и по другим причинам:

- они могут быть полезны с точными данными или без точных данных, а также любые данные требуют минимальной подготовки;
- новые опции могут быть добавлены к существующим деревьям;
- их ценность в выборе лучшего из нескольких вариантов;
- они легко сочетаются с другими инструментами принятия решений.

Однако деревья решений могут стать чрезмерно сложными. В таких случаях более компактная диаграмма влияния может быть хорошей альтернативой. Диаграммы влияния сужают акцент на критические решения, исходные данные и цели. Деревья решений играют существенную роль в машинном обучении и интеллектуальном анализе данных. Дерево решений также можно использовать для создания автоматизированных прогностических моделей, в которых используются приложения для машинного обучения, интеллектуального анализа данных и статистики. Этот метод, известный как изучение дерева решений, учитывает наблюдения в отношении элемента, чтобы определить его ценность.

Когда узлы представляют данные, а не решения, этот тип дерева используют как дерево классификации. Каждая ветвь содержит набор атрибутов или правил классификации, которые связаны с определенной меткой класса, которая находится в конце ветки. Эти правила, также известные как правила принятия решений, могут быть выражены в предложении if-then. Каждый дополнительный набор данных помогает модели более точно предсказать, какому из конечного набора значений принадлежит рассматриваемый объект. Затем эту информацию можно использовать в качестве входных данных в более широкой модели принятия решений. Иногда прогнозируемой переменной будет действительное число, например, цена. Деревья решений с непрерывными, вещественными результатами называются деревьями регрессии.

Для повышения точности иногда несколько деревьев используются вместе в методах ансамбля:

- создание пакетов предполагает получение нескольких деревьев путем повторной выборки исходных данных, после чего эти деревья используют для достижения консенсуса;
- классификатор случайного леса состоит из нескольких деревьев, предназначенных для увеличения скорости классификации;
- повышенные деревья, которые можно использовать для деревьев регрессии и классификации;

– все деревья в лесу вращения обучаются с использованием метода главных компонент на случайной части данных.

Дерево решений считается оптимальным, когда оно представляет наибольшее количество данных с наименьшим количеством уровней или вопросов. Алгоритмы, предназначенные для создания оптимизированных деревьев решений, включают CART, ASSISTANT, CLS и ID3/4/5. Дерево решений также можно создать, создав правила ассоциации, поместив целевую переменную справа. В каждом методе нужно определить, где лучше сделать разделение и на каком уровне. Обычные методы для этого включают применение коэффициента Джини, извлечение информации и уменьшение дисперсии. Использование деревьев решений в машинном обучении имеет несколько преимуществ:

- стоимость использования дерева для прогнозирования данных уменьшается с каждой дополнительной точкой данных;
- работает для категориальных или числовых данных;
- может моделировать проблемы с несколькими выходами;
- использует модель белого ящика (облегчая объяснение результатов) ;
- надежность дерева может быть проверена и определена количественно;
- имеет тенденцию быть точным независимо от того, нарушает ли оно допущения исходных данных.

Но у них также есть несколько недостатков:

- при работе с категориальными данными с несколькими уровнями прирост информации смещается в пользу атрибутов с наибольшим количеством уровней;
- расчеты могут стать сложными при работе с неопределенностью и множеством связанных результатов.

В появившихся исторически позднее древовидных картах отсутствует ввод вероятности сценария (выбора ветви) как первичное данное, но по размеру и цвету прямоугольника группы исходных данных можно получить наглядное представление об этой величине. Древовидные карты могут предоставлять существенные преимущества по сравнению с табличными данными в корпоративных приложениях благодаря присущей им способности поддерживать интеграцию пользователями разных измерений информации. Представления иерархии древовидных карт значительно различались по поддержке задач идентификации, сравнения и анализа, но были значительно быстрее и точнее, чем представления табличных данных. Изучение древовидной карты было, по крайней мере, таким же успешным, как и для таблиц. Различия в производительности между древовидными картами и таблицами увеличиваются с увеличением размера наборов данных. Субъективные оценки пользователей в подавляющем большинстве случаев поддерживали древовидные карты в табличных представлениях данных. Эти результаты показывают, что древовидные карты должны быть включены в качестве стандартного графического компонента в приложения для анализа и мониторинга данных на уровне предприятия.

Заключение

Тенденции развития древовидных карт являются частью общего процесса развития систем поддержки принятия решений, и достижения в этой области требуют комплексного изучения. Например, сейчас искусственный интеллект позволил людям принимать лучшие решения с помощью интеллектуальных DSS. Новые исследования, регулярно публикуемые по результатам конференций IFIP TC8 / Working Group 8.3 [11], показывают, что лица, принимающие решения, могут действовать более своевременно, используя данные в режиме реального времени, более точно благодаря интеллектуальному анализу данных и методам «больших данных», более стратегически, учитывая большее количество факторов, более инклюзивно, благодаря доступности данных социальных сетей и более широкий доступ к средствам массовой информации с помощью медиатехнологий. Начиная с 2014 г. [13] встречаются утверждения о рождении нового поколения технологий, которому присвоено название DSS2.0.

В рамках общего направления исследований формулируемые темы включают теоретические, эмпирические и проектные исследования; тематические подходы в системах поддержки принятия решений; модели принятия решений в реальном мире; информационные технологии в здравоохранении; теорию принятия решений; управление знаниями; открытие знаний и ресурсов; бизнес-аналитику; системы поддержки принятия групповых решений; «большие данные»; специальные

языки для поддержки принятия решений; мультимедийные инструменты для DSS; системы Web 2.0 в поддержке принятия решений и контекстные технологии для принятия решений; интеллектуальные системы и технологии поддержки принятия решений; методы исследования в DSS 2.0; мобильные DSS; аналитику социальных сетей. Для всех, кто разрабатывает или использует системы поддержки принятия решений, перечисленный список методов и приложений должен дать идеи и указания для вывода собственных разработок на новый технологический уровень.

Список литературы

1. Power D. J., Heavin C. *Decision Support, Analytics, and Business Intelligence*. Third Edition. New York : Business Expert Press, 2017. P. 196
2. Белозерцев А. И., Михеев М. Ю., Прокофьев О. В., Семочкина И. Ю. Методы анализа данных и их реализация в системах поддержки принятия решений. Аспекты инжиниринга информационно-измерительных систем : учеб. пособие. Пенза : Изд-во ПГУ, 2018. 148 с.
3. Прокофьев О. В., Семочкина И. Ю. Методы и средства поддержки принятия решений в социально-экономических задачах. Пенза : Изд-во ПГУ, 2010. 100 с.
4. Долгова И. А., Прокофьев О. В. Компьютерная поддержка принятия решений для управления социально-экономическими системами. Пенза : Изд-во ПГУ, 2012. 116 с.
5. Михеев М. Ю., Прокофьев О. В., Семочкина И. Ю. Методы анализа данных и их реализация в системах поддержки принятия решений : учеб. пособие. – Пенза : Изд-во ПГУ, 2014. 118 с.
6. Прокофьев О. В. Концепция и способы построения систем поддержки принятия решений на основе интеллектуального анализа данных // Системы интеллектуального анализа данных: методология, реализация, приложения. Пенза : АННОО ПДЗ, 2013. С. 22–40.
7. Тюрин М. В., Фомин С. А., Прокофьев О. В. Разработка моделей прогнозирования состояния элементов технически сложного объекта // Надежность и качество сложных систем. 2016. № 1. С. 73–78.
8. Николаев А. В., Прокофьев О. В., Тюрин М. В., Токарев А. Н. Большие данные для нового качественного уровня электроэнергетики // Труды Международного симпозиума Надежность и качество. 2018. Т. 2. С. 102–104.
9. Михеев М. Ю., Прокофьев О. В., Савочкин А. Е. Прогнозирование латентных ситуаций по коротким временным рядам // Труды Международного симпозиума Надежность и качество. 2017. Т. 1. С. 320–324.
10. What is a decision tree? URL: <https://www.lucidchart.com/pages/decision-tree>
11. Phillips-Wren G. E., Carlsson S., Respicio A. *DSS 2.0 – Supporting Decision Making With New Technologies*. IOS Press, 2014. P. 604.
12. Goldberg J. H., Helfman J. Enterprise Network Monitoring Using Treemaps. Human Factors and Ergonomics Society Annual Meeting Proceedings. September. 2005. P. 671–675. doi: 10.1177/154193120504900508.
13. International Federation For Information Processing. Technical Committee 8. Working Group 8.3. URL: <http://ifip8.dsi.uminho.pt/index.php/events?wg=8.3>

References

1. Power D.J., Heavin C. *Decision Support, Analytics, and Business Intelligence*. Third Edition. New York: Business Expert Press, 2017:196.
2. Belozertsev A.I., Mikheev M.Yu., Prokofev O.V., Semochkina I.Yu. *Metody analiza dannykh i ikh realizatsiya v sistemakh podderzhki prinyatiya resheniy. Aspekty inzhiniringa informatsionno-izmeritel'nykh sistem: ucheb. posobie = Methods of data analysis and their implementation in decision support systems. Aspects of information and measurement systems engineering : textbook*. Penza: Izd-vo PGU, 2018:148. (In Russ.)
3. Prokofev O.V., Semochkina I.Yu. *Metody i sredstva podderzhki prinyatiya resheniy v sotsial'no-ekonomicheskikh zadachakh = Methods and means of decision support in socio-economic problems*. Penza: Izd-vo PGU, 2010:100. (In Russ.)
4. Dolgova I.A., Prokofev O.V. *Komp'yuternaya podderzhka prinyatiya resheniy dlya upravleniya sotsial'no-ekonomicheskimi sistemami = Computer-based decision support for managing socio-economic systems*. Penza: Izd-vo PGU, 2012:116. (In Russ.)
5. Mikheev M.Yu., Prokofev O.V., Semochkina I.Yu. *Metody analiza dannykh i ikh realizatsiya v sistemakh podderzhki prinyatiya resheniy: ucheb. posobie = Methods of data analysis and their implementation in decision support systems: textbook*. Penza: Izd-vo PGU, 2014:118. (In Russ.)
6. Prokofev O.V. The concept and methods of building decision support systems based on data mining. *Sistemy intellektual'nogo analiza dannykh: metodologiya, realizatsiya, prilozheniya = Data mining systems: methodology, implementation, applications*. Penza: ANNOO PDZ, 2013:22–40. (In Russ.)

7. Tyurin M.V., Fomin S.A., Prokofev O.V. Development of models for predicting the state of elements of a technically complex object. *Nadezhnost' i kachestvo slozhnykh system* = Reliability and quality of complex systems. 2016;1:73–78. (In Russ.)
8. Nikolaev A.V., Prokofev O.V., Tyurin M.V., Tokarev A.N. Big data for a new qualitative level of the electric power industry. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo* = Proceedings of the International Symposium Reliability and Quality. 2018;2:102–104. (In Russ.)
9. Mikheev M.Yu., Prokofev O.V., Savochkin A.E. Predicting literal situations over short time series. *Trudy Mezhdunarodnogo simpoziuma Nadezhnost' i kachestvo* = Proceedings of the International Symposium Reliability and Quality. 2017;1:320–324. (In Russ.)
10. *What is a decision tree?* Available at: <https://www.lucidchart.com/pages/decision-tree>
11. Phillips-Wren G.E., Carlsson S., Respício A. *DSS 2.0 – Supporting Decision Making With New Technologies*. IOS Press, 2014:604.
12. Goldberg J.H., Helfman J. *Enterprise Network Monitoring Using Treemaps*. *Human Factors and Ergonomics Society Annual Meeting Proceedings*. September. 2005:671–675. doi: 10.1177/154193120504900508.
13. *International Federation For Information Processing. Technical Committee 8. Working Group 8.3*. Available at: <http://ifiptc8.dsi.uminho.pt/index.php/events?wg=8.3>

Информация об авторах / Information about the authors

Михаил Юрьевич Михеев

доктор технических наук, профессор,
заведующий кафедрой информационных
технологий и систем,
Пензенский государственный
технологический университет
(Россия, г. Пенза, проезд Байдукова/ул. Гагарина, 1а/11)
E-mail: mix1959@gmail.com

Олег Владимирович Прокофьев

кандидат технических наук, доцент
кафедры математики и информатики,
Пензенский государственный
технологический университет
(Россия, г. Пенза, проезд Байдукова/ул. Гагарина, 1а/11)
E-mail: prokof_ow@mail.ru.

Ирина Юриевна Семочкина

кандидат технических наук, доцент,
доцент кафедры информационных технологий и систем,
Пензенский государственный
технологический университет
(Россия, г. Пенза, проезд Байдукова/ул. Гагарина, 1а/11)
E-mail: ius1961@gmail.com

Mikhail Yu. Mikheev

Doctor of technical sciences, professor,
head of sub-department of informational technologies
and systems,
Penza State Technological University
(1a / 11 Baydukov passage /Gagarin street,
Penza, Russia)

Oleg V. Prokofev

Candidate of technical sciences, associate professor
of sub-department of mathematics and computer science,
Penza State Technological University
(1a / 11 Baydukov passage /Gagarin street,
Penza, Russia)

Irina Yu. Semochkina

Candidate of technical sciences, associate professor,
associate professor of sub-department of informational
technologies and systems,
Penza State Technological University
(1a / 11 Baydukov passage /Gagarin street,
Penza, Russia)