

Реферати

УДК 001.103:004.056

Ланде Д.В., Березін Б.О., Павленко О.Ю. Алгоритм сегментації слів на основі пошуку найкоротшого шляху в графі. *Регістрація, зберігання і оброб. даних*. 2017. Т. 19. № 4. С. 3–15. — рос.

Розглянуто особливості алгоритмів сегментації слів з текстів, які не містять роздільників. Представлено новий алгоритм сегментації слів на основі пошуку найкоротшого шляху. Наведено оцінки якості сегментації. Показано можливості використання наведеного алгоритму в задачах пошуку інформації у національних доменах мережі Інтернет. Реалізацію алгоритму сегментації слів використано для створення узагальненої моделі предметної області на базі моніторингу ресурсів китайського сегменту Інтернет. Табл. 3. Іл.: 5. Бібліогр.: 23 найм.

Ключові слова: сегментація слів, сегментатор, якість сегментації слів, моніторинг, пошук найкоротшого шляху, хвильовий алгоритм.

UDC 001.103:004.056

Lande D.V., Berezin B.A. and Pavlenko O.Yu. Algorithm for word segmentation based on the shortest path algorithm inside a graph. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 3–15. — Rus.

The features of word segmentation algorithms from such texts are considered. There are two main models, namely, statistical one and the one using a dictionary. For models with a dictionary, a variant of the maximal matching algorithm is noted for which there are modifications such as Forward Maximal Matching (FMM) and Backward Maximal Matching (BMM) to be depending on the direction of text processing. The second option for models with a dictionary is an algorithm that finds segmentation with a minimum number of words. A new algorithm for words segmenting based on a modified wave algorithm has been presented. The algorithm takes into account the features of the input data and is built in such a way that the necessary calculations are performed in a single pass. This reduces its computational complexity. A description of the word segmentation algorithm is given. An example is shown of splitting an input string in English into words, representing it in the form of a graph and finding the shortest path.

To assess the quality of segmentation, the EDWS (Edit Distance of the Word Separator) method is presented. A special tool was used to assess the segmentation of Chinese words with a test corpora based on news texts. Evaluations of the quality of segmentation of words for the proposed algorithm (based on the search for the shortest path) and a number of other known segmentators are obtained. An example of segmentation of a news text in Russian is given. The possibilities of using the developed algorithm in the problems of information search in national resources of the Internet are shown. The implementation of the word segmentation algorithm is used when creating a generalized domain model based on monitoring of the Chinese Internet segment resources.

The increase in the number of information resources of the Chinese Internet segment makes it necessary to create of global information retrieval systems. For search indexes of such systems, fast, accurate and complete segmentation of words from texts is necessary. The obtained estimates of segmentation quality using the proposed algorithm for the formation of the search system index indicate the possibility of its use for information resources of the Chinese Internet-segment. Tabl.: 3. Fig.: 5. Refs: 23 titles.

Key words: words segmentation, word segmentation quality, web resources monitoring, shortest path search, wave algorithm.

УДК 004.942.519.87

Додонов А.Г., Додонов В.А., Кузьмичов А.И. Визуальная поддержка оптимальных решений в пространственных сетях. *Регистрация, хранение и оброб. данных*. 2017. Т. 19. № 4. С. 16–25. — укр.

Практика оптимізаційного моделювання задач комунікації в просторових мережах стикається з необхідністю оперативного отримання наряду з знайденим оптимальним планом відповідного зображення. Предложена композиція доступних програмних продуктів, використовуваних на звичайних ПК (АРМ), її використання дозволяє отримувати зважені управлінські рішення в режимі динамічного змінення значень наборів даних, користуючись побудованим наглядним графом по результатам оптимізаційних обчислень. Ил.: 6. Библиогр.: 11 найм.

Ключевые слова: просторова мережа, задача комівояжера, візуалізація графів, сетеве моделювання задач оптимізації, генетичний алгоритм, Excel Solver, yEd Graph Editor.

UDC 004.942.519.87

Dodonov A.G., Dodonov V.A. and Kuzmichov A.I. Visual support of optimal decisions in spatial networks. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 16–25. — Ukr.

The investigation deals with complex application of graph visualization and business analytics for obtaining solutions of better quality, which is difficult to do utilizing traditional means, so graphs have played a role in many of those solutions. The function of a graph is to represent links between things such as a spatial network, revealing the organization structure. Relationships are fundamental for the spatial network, which is one of the reasons of graph analysis and visualization having so much potential for value. Nowadays, several of our most significant research and software development efforts are, in essence, graph-based. The graph-based approaches gain deeper understanding of the dynamics of organizational and business processes.

The practice of optimization simulation of communication problems in spatial networks is faced with the need to quickly obtain the corresponding image with the optimal plan. The use of computer-rendered visualization for decision-making in business is a relatively recent phenomenon. Thus, the composition of available graph visualization software used by ordinary PCs is offered. Its usage allows obtaining weighted management decisions for dynamically changing values of data sets using the constructed graph on the results of optimization solutions. Fig.: 6. Refs: 11 titles.

Key words: spatial network, traveling salesman problem, graph visualization, network modeling of optimization problems, genetic algorithm, Excel Solver, yEd Graph Editor.

УДК 551.508.85 (551.576 +551.577)

Путятін В.Г., Корбан Д.В., Князь О.І. Вплив опадів на роздільну здатність суднової радіолокаційної станції за кутовими координатами. *Регістрація, зберігання і оброб. даних*. 2017. Т. 19. № 4. С. 26–34. — рос.

Проведено аналіз впливу випадного дощу на роздільну здатність суднової радіолокаційної станції (РЛС) за кутовими координатами, яка працює в міліметровому та сантиметровому діапазонах довжин хвиль. Показано, що в сантиметровому діапазоні довжин хвиль ($\lambda = 3 \text{ см} \dots 10 \text{ см}$) має місце поліпшення роздільної здатності суднової РЛС за кутовими координатами, а в міліметровому діапазоні довжин хвиль має місце як поліпшення роздільної здатності суднової РЛС за кутовими координатами, так і її погіршення. Библиогр.: 11 найм.

Ключові слова: кутові координати, розсіювання, поглинання, випромінювання, функція розсіювання, роздільна здатність, діаграма спрямованості антени, межі розділу, тілесний кут, суднова радіолокаційна станція.

UDC 551.508.85 (551.576 +551.577)

Putyatin V.G., Korban D.V. and Knyaz A.I. Influence of precipitations on angular coordinates resolution for a radar. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 26–34. — Rus.

An analysis of influence of rain drop down on the resolution power of marine radar by angular coordinates, working in the millimeter and centimeter wavelength ranges has been performed. It is shown that one of the directions of increase of safety of navigation using marine radar has the improvement of their resolution in angular coordinates, since the atmospheric phenomenon on the path of the ship make a significant error in the measurement of the angular coordinates by the radar surveillance navigation facilities. The study of precipitation on the resolution power of marine radar by angular coordinates, antenna which radiates a linear-polarized electromagnetic wave of millimeter and centimeter ranges has been provided. The interaction of the wave with particles of precipitation presented in the form of a power wave from the effective surface of the scattering reflective volume of rain is taken into account. The effective surface of scattering and attenuation on all particle sizes in the dropdown of rain on the way is found. It was found that for the centimeter wavelength range ($\lambda = 3 \text{ cm} \dots 10 \text{ cm}$) attenuation is much higher than the scattering. The resulting conditions to improve the resolution of the antenna of marine radar by angular coordinates are in the centimeter wavelength range. In the millimeter range of wavelengths occurs as the improvement of the resolution capability of the antenna of marine radar and its deterioration depending on the intensity of precipitation. The deterioration of the resolution can reach 10 %. Refs: 11 titles.

Key words: angular coordinates, dispersion, absorption, radiation, function of dispersion, resolution, diagram of orientation of aerial, border of division, corporal corner, ship radar.

УДК 004.942

Кузнецова Н.В. Разработка скоринговых карт для анализа рисков банковской деятельности. *Регистрация, хранение и обраб. данных*. 2017. Т. 19. № 4. С. 35–44. — укр.

Статья посвящена анализу особенностей разработки скоринговых карт: последовательности этапов, действующих лиц, методологии разработки. На примере анализа кредитных рисков показано, как формируется скоринговая карта и рассчитываются скоринговые баллы, и каким образом учитывается профиль риска в скоринговой карте. Показано, что скоринговые карты являются эффективным инструментом в виде табличных оценок для работы банковских кредитных аналитиков. Использование скоринговых карт позволяет оперативно учитывать изменения на рынке банковских услуг, оперируя разными порогами отсечения и оценивая, какую степень риска на себя принимает банк. Табл.: 1. Ил.: 3. Библиогр.: 6 наим.

Ключевые слова: скоринговые карты, кредитные риски, профиль риска, скоринговые баллы, толерантность к риску.

UDC 004.942

Kuznietsova N.V. Scoring card's development for banking activities risks' analysis. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 35–44. — Ukr.

The basic features of scoring card building such as a sequence of stages, roles and methodology of development are considered. On example of the credit risks analysis it is shown how scoring card is built and scoring points are calculated. It is also shown how by varying of the threshold to receive different rate of good and bad credit applications which would be accepted. Also scorecard gives the possibility to define the size of risk by bad rate of credit application on each interval of the scorecard points. Moreover it is recommended, in scoring cards, to take into account the risk's profile that means to pay attention not only on demographic characteristics of the applicants but also on their behaviour in the past. It is demonstrated that the scoring cards are effective instrument in table view of scoring points for bank credit analytics. Using the scoring cards allows taking into account changes in the banking services market, operating with different cut-off thresholds and evaluating which degree of risk the bank takes on itself. Tabl.: 1. Fig.: 3. Refs: 6 titles.

Key words: scoring cards, credit risks, risk's profile, scoring scores, risk tolerance.

УДК 004.056.2

Василенко В.С., Матов А.Я. Методика определения выходных данных для оценки остаточных рисков при обеспечении конфиденциальности информационных объектов. *Регистрация, хранение и обраб. данных*. 2017. Т. 19. № 4. С. 45–55. — укр.

Для анализа защищенности информации автоматизированных систем предложено применение количественных характеристик в виде величин остаточного риска или вероятности преодоления нарушителем средств защиты тех или других свойств защищенности. Для оценки защищенности конфиденциальной информации телекоммуникационных сетей с использованием соответствующих систем защиты информации предложены математические модели для определения вероятностей преодоления нарушителем составляющих системы защиты конфиденциальности. Табл.: 1. Ил.: 1. Библиогр.: 7 наим.

Ключевые слова: угроза, защита, конфиденциальность, модель, нарушитель, ресурсы, телекоммуникационная сеть.

UDC 004.056.2

Vasilenko V.S. and Matov O.Ya. Methods on determination of the output data for assessing residual risks with ensuring the confidentiality of information objects. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 45–55. — Ukr.

An application of quantitative characteristics in the form of values of residual risk or probabilities of overcoming by the violator of protection facilities of different security properties is offered for the analysis of information security of the automated systems. To assess the protection of confidentiality of telecommunications networks using appropriate information security systems, a mathematical model for to determine the probability of overcoming the offender components of privacy is proposed. Tabl.: 1. Fig.: 1. Refs: 7 titles.

Key words: threat protection, confidentiality, model, the offender, resources, telecommunications network.

УДК 004.415.26:004.415.2

Гайдаржи В.И., Савченко Н.Н. Средства визуализации классов и взаимосвязей объектно-реляционной базы данных Caché. *Регистрация, хранение и обраб. данных*. 2017. Т. 19. № 4. С. 56–63. — укр.

Представлен анализ средств визуализации состава классов, определенных в объектно-реляционной базе данных Caché корпорации InterSystems. Проведено сравнение разных подходов к отображению классов Caché и определены недостатки существующих средств визуализации. Кроме того, выявлена определенная недостаточность традиционной UML-диаграммы классов для представления составляющих классов Caché. Предложено соответствующее расширение нотации UML и способ для наиболее информативной визуализации классов Caché, который обеспечивает отображение элементов, характерных классам БД Caché. Приведено описание веб-интерфейса визуализации. Ил.: 5. Библиогр.: 6 наим.

Ключевые слова: визуализация, InterSystems Caché, диаграмма классов.

UDC 004.415.26:004.415.2

Gaidarzhy V.I. and Savchenko M.M. Methods for visualization of classes and interconnections of object-relational database Caché. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 56–63. — Ukr.

The classes and their relationships in hybrid, post-relational (object-relational) database management system InterSystems Caché (InterSystems IRIS platform) are analyzed. A better notation to visualize that class entities and relationships is proposed. The different approaches to visualize the Caché classes and identify limitations of existing visualization tools are compared. Due to the fact that there are

many features in InterSystems Caché classes and relationships, a new notation development is required to model the existing systems with all of their features displayed. Authors propose to extend UML notation and compare it with existing UML solutions, which are not fully capable of visualizing all features under Caché DBMS. This article also provides An information over a developed standard extension and a web-based software which implements this standard is also provided. The features developed web-based solution for visualizing classes and their relationships, comparing it to existing visualization tools are described. Fig.: 5. Refs: 6 titles

Key words: visualization, InterSystems Caché, class diagram.

УДК 004.5

Бойченко А.В., Ландэ Д.В. Формирование сценариев управленческой деятельности на основе изучения информационного пространства. *Регистрация, хранение и обраб. данных*. 2017. Т. 19. № 4. С. 64–70. — укр.

Рассмотрены технологические решения, позволяющие анализировать массивы полнотекстовых документов, и предложена процедура автоматизированного построения сценариев управленческой деятельности на основе анализа информационного пространства для выбранных предметных областей. Ил.: 3. Библиогр.: 8 наим.

Ключевые слова: сценарный анализ, управленческая деятельность, информационные технологии, семантическая сеть.

UDC 004.5

Boychenko A.V. and Lande D.V. Formation of management scenarios by means of investigation of information space. *Data Rec., Storage & Processing*. 2017. Vol. 19. N 4. P. 64–70. — Ukr.

It is considered the development of a scenario approach and the construction of an information technology, which allows in an automated mode, based on the analysis of the input documentation package, to solve the problem of the formation and ranking of scenarios of influencing the objects that correspond to the selected key concepts.

A mathematical model of qualitative estimation of management activity scenarios for a subject area is worked out. To construct a scenario, an abstract automaton is proposed that describes the requirements based on the description of the analytic task. The scenarios are singled out from the semantic network. There are the vertices of concepts, and the arcs to be mutual relations, for example, «cause – consequence». A modified ant colony optimization (ACO) approach for forming scenarios is proposed. To check out the suggested technology, text domains were created for such subject areas as «Brexit», «Mass Shuting», «Catalan referendum», «Saudi Arabian purge», «Zimbabwean coup d’etat», «Syrian civil war». Fig.: 3. Refs: 8 titles

Key words: scenario analysis, management activity, information technologies, semantic network.
