INTUITIVE CONTENT MANAGEMENT SYSTEM VIA MANIPULATION AND DUPLICATION WITH IF-ELSE RULES CLASSIFICATION

CHAN CHUNG HOONG

Master of Science

UNIVERSITI MALAYSIA PAHANG

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Master of Science

_____

(Supervisor's Signature)

Full Name : DR.MOHAMED ARIFF BIN AMEEDEEN

Position : SENIOR LECTURER

Date

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

_____

(Student's Signature)

Full Name :CHAN CHUNG HOONG

ID Number : MCC14005

Date :

INTUITIVE CONTENT MANAGEMENT SYSTEM VIA MANIPULATIONB AND DUPLICATION WITH IF-ELSE RULES CLASSIFICATION

CHAN CHUNG HOONG

Thesis submitted in fulfilment of the requirements
for the award of the degree of
Master of Science

Faculty of Computer System & Software Engineering

UNIVERSITI MALAYSIA PAHANG

SEPTEMBER 2018

ACKNOWLEDGEMENTS

I would like to thank the society for their participation in the survey who supported my work in this way and helped me get results of better quality. A very special gratitude goes out to all down at Research Fund and also Rich Foundation for helping and providing the funding for the work.

I would like to thank my supervisor, Dr. Mohamed Ariff Bin Ameedeen for his invaluable advice and contributions to this work. Her insights and high standards have definitely helped to shape this work. It is a pleasure to have an advisor being so joyful in his work.

I would like to thank my friends for accepting nothing less than excellence from me. Last but not the least, I would like to thank my family: my parents and to my brothers and sister for supporting me spiritually throughout writing this thesis and my life in general.

This project is successfully completed with the assistance of Jabatan Pembangunan Wanita (JPW). I am thankful for this project with its grant by JPW for helping rural areas to generate website for their businesses.

Last but not least, I would like to express my appreciation towards FSKKP for their assistance throughout this project. With all assistances from all parties above, my project was successfully completed and presented.

ABSTRACT

A Content Management System (CMS) is a system that is used to control or manage content on a website such as text, links, image, HTML documents, and other forms of media. The fact that CMS is dynamically generated is a problem on pages loading speed which is currently used in Drupal, Joomla, WordPress, and Agility CMS. However, static websites load faster than dynamic ones, since the content is delivered as-is and can be cached by a content delivery network (CDN), and the web server does not need to perform any application logic or database queries. In order to solve this problem, an Intuitive Content Management System (ICMS) is proposed in this study. The aim of ICMS is to help non-technical users to design their website easily. Users will need to choose the content from templates, upload images directly, and fill in necessary documents to create a quick, easy, and static website which can reduce load time from dynamic website. An ICMS is a content management system that uses three main important techniques to build, such as Manipulation method, Data Mining Prediction, and Boyer-Moore Horspool algorithm. The manipulation method is a method that can easily generate a website and store the information precisely in a correct order with buffering method. A buffering method will be implemented after the generation of website, in order to transform dynamic website into static website. Therefore, a static website is created which load faster than dynamic website. An ICMS needs to use data mining classification to accurately classify stored data while predicting the user's database information to give the best preferences and suitable template chosen for an end user to select. The time consumed for choosing a template will decrease as an example of giving the user 5 templates to choose from out of 100. As a result, it can minimize the workload of choosing a template and the time consumed which come from preferences of 5 recommendation which are 20 times faster than before. Boyer-Moore Horspool algorithm is adapted in ICMS which is mainly for faster searching result. As a result, ICMS can transform dynamic websites into static websites with faster load speed using manipulation method mixed with data mining classification prediction and Boyer-Moore Horspool algorithm which can be classified, edited, adjustable and searched more precisely. An ICMS makes things easy for any small business or personal business to create, manage, and publish interactive observations to interpret their own websites, which also support mobile platform browsers.

ABSTRAK

Sistem Pengurusan Kandungan (CMS) merupakan satu sistem yang digunakan untuk mengawal atau mengurus kandungan di laman web seperti teks, pautan, imej, dokumen HTML, dan bentuk media lain. Apabila CMS dihasilkan secara dinamik, ia merupakan satu masalah dalam kelajuan memuat turun laman web seperti Drupal, Joomla, WordPress, dan Agility CMS. Walau bagaimanapun, laman web statik memuat lebih pantas daripada yang dinamik, ia kerana kandungan boleh dihantarkan sebagai cache dalam bentuk content delivery network (CDN). Ini boleh menyebabkan laman web tidak perlu melakukan sebarang pertanyaan logik aplikasi atau pangkalan data. Sebagai penyelesaian masalah ini, Sistem Pengurusan Kandungan Secara Intuitif (ICMS) dicadangkan dalam kajian ini. Tujuan ICMS ini adalah untuk membantu pengguna bukan teknikal untuk mereka bentuk laman web mereka dengan lebih mudah. Pengguna-pengguna hanya perlu memilih kandungan dari templat, memuat naikan gambar secara langsung, dan mengisi dokumen yang diperlukan untuk membuat laman web yang cepat, mudah, dan secara statik boleh mengurangkan masa memuat turun lebih pantas dari laman web dinamik. ICMS adalah sistem pengurusan kandungan yang menggunakan tiga teknik utama seperti kaedah Manipulasi, Prediksi Data Mining, dan algoritma Boyer-Moore Horspool. Kaedah manipulasi adalah satu kaedah yang dapat menjana laman web dan menyimpan maklumat dengan tepat dalam susunan yang betul dengan kaedah penimbunan. Kaedah penimbunan akan dilaksanakan selepas penubuhan laman web, untuk pengubahan laman web dinamik kepada laman web statik. Oleh itu, laman web statik melayari lebih cepat daripada laman web dinamik. ICMS perlu menggunakan klasifikasi Data Mining untuk mengklasifikasikan data yang disimpan dengan tepat dan meramalkan maklumat pangkalan data pengguna untuk memberikan pilihan terbaik dan templat yang sesuai dipilih untuk pengguna pilihan. Masa yang digunakan untuk memilih templat akan berkurangan. Sebagai contoh, ia memberi pengguna lima template untuk dipilih dari seratus. Hasilnya, ia dapat mengurangkan beban kerja memilih templat dan masa yang digunakan dari pilihan 5 cadangan menimbul 20 kali kanda lebih pantas dari sebelumnya. Algoritma Boyer-Moore Horspool disesuaikan dalam ICMS yang terutama untuk menghasilkan pencarian lebih pantas. Akibatnya, ICMS boleh mengubah laman web dinamik ke laman web statik dengan kelajuan beban lebih cepat dalam penggunaan kaedah manipulasi yang bercampur dengan ramalan klasifikasi data mining dan algoritma Boyer-Moore Horspool yang boleh diklasifikasikan, diubahsuai, diselaraskan dan carian lebih tepat. ICMS memudahkan semua perniagaan kecil atau perniagaan peribadi untuk membuat, mengurus, dan menerbitkan pemerhatian interaktif dalam pentafsir laman web mereka sendiri, dan juga menyokong pelayar platform lebih mudah alih.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CMS	Content Management System
ICMS	Intuitive Content Management System
PHP	Hypertext Pre-processor
CSS	Cascading Style Sheets
HTML	Hypertext Mark-up Language
XHTML	Extensible Hyper Text Markup Language
URL	Universal Resource Locator
MYSQL	My Structured Query Language
POSTGRESQL	Postgres Structured Query Language
SAAS	Software as a Service

REFERENCES

Abdulwahid, A. A. (2015) Drupal Content Management System.

Arumugam, J. (2016). Drupal: An Emerging Content Management System For Libraries. *Journal of Library, Information and Communication Technology*, *7*(1-2), 11.

Awad, H. A. (2014). Cloud Computing as an Operational Model for ERP Services: Definitions and Challenges. *International Journal of Innovation and Applied Studies*, *8*(2), 499.

Boyer, R. S., & Moore, J. S. (1977). A Fast String Searching Algorithm. *Communications of the ACM*, *20*(10), 762-772.

Chakrabarti, S., Ester, M., Fayyad, U., Gehrke, J., Han, J., Morishita, S., ... & Wang, W. (2006). Data Mining Curriculum: A Proposal (Version 1.0). *Intensive Working Group of ACM SIGKDD Curriculum Committee*, 140.

Chen, J., Cai, S., Zhu, L., Guo, Y., Huang, R., Zhao, X., & Sheng Y.,(2016). An Improved String-Searching Algorithm and Its Application In Component Security Testing. *Tsinghua Science and* Technology, *21*(3), 281-294.

Clifton, C. (2017). Encyclopædia britannica: Definition of Data Mining. *Britannica.com.* Retrieved from https://www.britannica.com/technology/data-mining.

Crashin, S., Usoni, R., & Chandler, A. (2015). Data Mining Methods: Strategies and Algorithms on Different Applications. *Journal of Mechatronics*, *3*(1), 37-48.

Dorigo, M., Maniezzom V., and Colorni, A. (1996). Ant System: Optimization by a Colony of Cooperating Agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 26*(1), 29-41.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, *17*(3), 37.

Fitzpatrick, C. (2010). Improving Product-Market Fit by Engaging an Open Source Community. *Open Source Business Resource*, *July 2010*.

George, N. (2015). Beginning Django CMS: Your Blog Website: Templates, *Apress*, 45-49.

Han, J., & Kamber, M. (2000). Data Mining: Concepts and Techniques. *The Morgan Kaufmann Series in data management systems*, 327-400.

Han, J., Pei, J., & Kamber, M. (2011). Data Mining: Concepts and Techniques. *Elsevier*, 327-450.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised Learning: In The Elements of Statistical Learning. *Springer New York*, 485-585.

Itani, W.; Kayssi, A.; Chehab, A., (2009) Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures in Dependable, Autonomic and Secure Computing. *Proceedings of the 8th IEEE International Conference on Dependable, Autonomic and Secure Computing, Dec 2009*, 711-716.

Jabeen, H., & Baig, A. R. (2013). Two-stage Learning for Multi-Class Classification Using Genetic Programming. *Neurocomputing, 116*, 311-316.

Jacobson, L., & Kanber, B. (2015). Introduction In Genetic Algorithms in Java Basics. *Apress*, 1-19.

John, S. (2014). 55+ Best Responsive Blogger Templates for 2015. *Designssave.com.* Retrieved from https://designssave.com/best-responsive-blogger-templates.html.

Karp, R. M., & Rabin, M. O. (1987). Efficient Randomized Pattern-Matching Algorithms. *IBM Journal of Research and Development*, *31*(2), 249-260.

Kozielski, M., Doetsch, P., & Ney, H. (2013). Improvements in RWTH's System for Off-line Handwriting Recognition. *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR), 2013*, 935-939.

Kriegel, Hans-Peter; Kröger, Peer; Sander, Jörg & Zimek, Arthur (2011). Density-Based Clustering. *WIREs Data Mining and Knowledge Discovery. 1*(3): 231–240.

Kulekci, M. O. (2007). Tara: An Algorithm for Fast Searching of Multiple Patterns on Text Files. *Computer and Information Sciences, 2007, 22nd International Symposium on IEEE*, 1-6.

Langdon, W. B., & Poli, R. (2013). Foundations of Genetic Programming. *Springer Science & Business Media*.

Larkin, M. (2014). Getting Started with Shopify. Shopify Application Development Build Highly Effective Shopify Apps Using the Powerful Ruby on Rails Framework, 6.

Larose, D. T. (2014). Discovering Knowledge in Data: An Introduction to Data Mining. *John Wiley & Sons*, 10-15.

Lynn Beighley, Seamus Bellamy (2011). Drupal For Dummies, *2nd Edition*, *02*, 5-8.

Mandumula, K. K. (2011). Knuth-Morris-Pratt Algorithm. *Poslední zmena*, 18.

Martynov, M., & Novikov, B. (1996). An Indexing Algorithm for Text Retrieval. *ADBIS*, 171-175.

Moh'd Mhashi, M., & Alwakeel, M. (2010). New Enhanced Exact String Searching Algorithm. *Proceeding of the 1st International Journal of Computer Science and Network Security* (*IJCSNS), 10*(4), 193.

Mohammad, A., Saleh, O., & Abdeen, R. A. (2006). Occurrences Algorithm for String Searching Based on Brute-force Algorithm. *Journal of Computer Science, 2*(1), 82-85.

Nathalie Japkowicz and Mohak Shah. (2011). Evaluating Learning Algorithms: A Classification Perspective. *Cambridge University Press, New York, NY, USA*, 191-195.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A Method for Taxonomy Development and its Application in Information Systems. *European Journal of Information Systems, 22*(3), 336-359.

Niraj Singhal, Tanmeya Mohan, Subham Sarkar (2009). A Comparative Study Based On Open Source Content Management Systems. *Indian Journal of Computer Science and Engineering, 1*(4), 267-276.

Niraj Singhal, Tanmeya Mohan, Subham Sarkar (2009). A Comparative Study Based On Open Source Content Management Systems. *Indian Journal of Computer Science and Engineering, 1*(4), 267-276.

Nood.org Team (2009). Basic Content Management in Drupal. *Nood.org Team*. Retrieved from http://nood.org/sites/noodorg/files/page/basic-content-management-in-drupal-noodorg.pdf, 1-15.

Omar, M., Soliman, M., Ghany, A. A., & Bendary, F. (2013). Optimal Tuning of PID Controllers for Hydrothermal Load Frequency Control Using Ant Colony Optimization. *International Journal on Electrical Engineering and Informatics*, *5*(3), 348.

Panda, M. K. (2013, May). Empirical Testing Of The Neural Network Application Using Feedforward Testing Method. *International Journal of Engineering Research and Technology*, ESRSA Publications, *2*(6), 163-175.

Patel, S.K.; Rathod, V.R.; Parikh, S., (2011). Joomla, Drupal and WordPress - a Statistical Comparison of Open Source CMS. *Trendz in Information Sciences and Computing (TISC), 2011 3rd International Conference*, 182-187.

Priefer, D. (2014). Model-Driven Development of Content Management Systems Based on Joomla. *Proceedings of the 29th ACM/IEEE International Conference on Automated Software Engineering*, *ACM*, 911-914.

Quadri, S. A. (2011). Developing, Managing and Maintaining Web Applications with Content Management Systems: Drupal and Joomla as Case Study. *HAAGA-HELIA University of Applied Sciences*, 12-15.

Quinlan, J. R. (2014). C4.5: Programs for Machine Learning. *Morgan Kaufmann*, 200.

Reese, G. (2008). Cloud Application Architectures: Building Applications and Infrastructure in the Cloud. *Cloud Computing, 1*(1), 1-2.

Rokach, L. & Oded, M. (2005). Clustering Methods: Data Mining and Knowledge Discovery Handbook. *Springer US*, 321-352.

Saha, S., & Bandyopadhyay, S. (2013). A Generalized Automatic Clustering Algorithm in a Multiobjective Framework. *Applied Soft Computing, 13*(1), 89-108.

Sakusa, T., & Toyama, M. (2015). Supporting Web Content Development Using Web Index. *Proceedings of the 19th International Database Engineering & Applications Symposium*, *ACM*, 204-205.

Seifert, J. W. (2004). Data Mining: An Overview. *CRS Report for Congress. Order Code RL31798*, 3-12.

Sfetcu, N. (2015). Small Business Management for Online Business. *Raleigh, North Carolina: Lulu Press, Inc*, 192-198.

Siong, K. C.., (2015) What Are the Advantage/Disadvantages of Using Shopify to Build an E-commerce Website? *Quora.com,* Retrieved from https://www.quora.com/What-are-the-advantage-disadvantages-of-using-Shopify-to-build-an-e-commerce-website.

Smith, J. W., & McGuffee, J. W. (2015). Data and Content Management: A Master's Level Course. *Journal of Computing Sciences in Colleges, 31*(1), 70-79.

Squier, L. (2001). What is Data Mining? Reston, VA: Data Management Association National Capital Region, 7-12.

Sun, R. (2004). Desiderata for Cognitive Architectures. *Philos, Psychol. 17*, 341–373.

Tummalapalli, S., & Rao Machavarapu, V. (2016). Managing Mysql Cluster Data Using Cloudera Impala. *Procedia Computer Science, 85*, 463-474.

Vazirani, V. V. (2013). Approximation Algorithms. *Springer Science & Business Media,* 77-86.

Wakode, B.V. & Chaudhari, D.N. (2013). Study of Content Management Systems Joomla and Drupal. *International Journal of Research in Engineering and Technology*, *2*(12), 7.

Web Services, Amazon. (2017). Hosting Static Websites on AWS. *Awsstatic.com*. Retrieved from https://d0.awsstatic.com/whitepapers/Building%20Static%20Websites%20on%20AWS.pdf

Xhaferi, G., Memeti, A., & Imeri, F. (2015). Comparison of Several Algorithms for Searching Data's in a Learning Management System. *Proceeding of the 4th Mediterranean Conference on 2015 IEEE Embedded Computing (MECO)*, 272-275.