Improving Premium Domain Names Registration for .PS Domain Cctld Basing on Knowledge Discovery and Data Mining Techniques

Ibrahim Alfayoumi¹, Wael Al Sarraj²
Faculty of Information Technology, Islamic University of Gaza,
Gaza, Palestine, me@ibrahim.ps, wsarraj@iugaza.edu.ps

Abstract—Country code top-level domain represents the identity of its own country. About ten million Palestinians spread in the world belong or affiliate to the domain name (.PS), for which it represents their identity although they do not have their own state. In addition, the (.PS) represents Palestinians history, culture, and lifestyle as well as to the Palestine homeland issue on the cyber and internet space. Domain names are valuable in today's e-commerce and are considered one of the potential intellectual capital assets worldwide. However, Premium Domain Names (PDN) is generic single keywords that are high value, memorable and easily marketable; its cost can be significantly more than a typical domain purchase due to its perceived higher value. In this paper, the research is based on Knowledge Discovery and Data Mining Techniques to discover the most attention regions in the world that register .PS domains in order to manage acquisition knowledge, which will benefit in providing marketing plans, identify stakeholders and target customers. Two phases were performed to achieve the required work plan: Data Mining (the knowledge discovery) phase and Knowledge Management (sharing and planning). The techniques in both phases were chosen to give the best accuracy and particularity of the involved target audience in the study. The results show that what we suggested could be efficiently used to recognize patterns in such data set behavior to generate new marketing strategies and plans.

Index Terms—DNS, Data Mining, DNS Mining, Classification, Knowledge Discovery, Domain Name Marketing.

I INTRODUCTION

In today's ecommerce, domain names are a valuable intellectual capital asset, that is domains are unique, represents a unique IP address or a location/address on the Internet, they often are comprised of words that introduce valuable trademarks and businesses spend huge amounts of money for marketing their domain names to facilitate access and traffic to their websites. Premium domain names prices are usually higher than any regular domain because of because of its highly marketing value. The importance of premium domains appears in all novel internet usages, such as internet of things (IoT), e-marketing and e-branding [1-3]. Country code top-level domain represents the identity of its own country. About ten million Palestinians spread in the world belong or affiliate to the domain name (.PS), for which it represents their identity although they do not have their own state. Also the .PS represents Palestinians history, culture, lifestyle and in addition to the Palestine homeland issue on the cyber and internet space. In addition to premium names, (.PS) domain is characterized by its letters which can be used to create distinctive names, for example (tri.ps) as a domain that could represent the travel and tourism companies, (ma.ps) which represents Maps and GPS locations and the domain (li.ps) as a Cosmetics companies. All of the above gives a clear impression on the importance of this domain name as an intellectual capital, could support the development of generating perfect marketing plans and strategies, and could make remarkable revenue. Data mining (DM) is a process of discovering patterns in data sets, a pattern is an arrangement of repeated parts, which represents knowledge, or we can say that it is a set of rows that share the same values in two or more columns of a data table. DM involves machine learning, artificial intelligence, statistics, and database systems. Extracting information from and transform it into knowledge as an understandable structure for further use is DM [4, 5]. DM also involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating as shown Figure 1. DM is the analysis step of the Knowledge Discovery in Databases process (KDD) [6, 7].

Figure 1: Data Mining Process [20].
Knowledge Management (KM) is a process of creating and utilizing knowledge[8]. KM processes as shown Figure 2 can be integrated with a corporation system which can help the marketers to get the knowledge easily [9]. Good marketing decisions are based on knowledge that comes from customer’s behavior. It’s considered as a key for the marketing functions and can be found in the organization’s databases but most of it is hidden [10]. Marketing decisions are very important for any organization to increase profit and can affect customer’s behavior.

II RELATED WORKS

A systematic methodology that uses data mining and knowledge management techniques was proposed by Shaw, M.J., et al. to manage the marketing knowledge and support marketing decisions. This methodology can be the basis for enhancing customer relationship management [8].

Shu-hsien Liao, Yin-ju Chen and Hsin-hua Hsieh analyzed consumer adumbration, lifestyle habits and purchasing behavior in an application of Internet marketing to the direct selling industry and the cosmetics market in Taiwan by implementing association rules and cluster analysis as approaches for data mining [9].

Sérgio Moro and Raul M. S. Laureano described an implementation of a DM project based on the CRISP-DM methodology. Their data set was collected from a Portuguese marketing campaign related with bank deposit subscription. Their business goal was to find a model that can explain success of a contact. They claimed that their model can increase campaign efficiency, helping in a better management of the available resources and selection of a high quality and affordable set of potential buying customers [10].

Social networks have generated great expectations connected with their potential business value. Surma, J. and A. Furmanek purposed a research to prove that data mining techniques can bring statistically significant improvement in marketing response accuracy throughout the virtual community. In their test (classification and regression tree) approach was used to generate a classification tree to formulate specific rules to identify the proper target group and showed that it is possible to improve marketing response [11].

In 1997, Karl M. Wiig clarified that, progressive managers consider intellectual capital management (ICM) and knowledge management (KM) to be vital for sustained viability. Recent practices support this notion and have provided important approaches and tools KM supports ICM by focusing on detailed systematic, explicit processes and overlap and synergy between ICM and KM. and advanced enterprises pursue deliberate strategies to coordinate and exploit them [12].

Ya-Hui Ling, selected samples from a list of the top 1,000 Taiwanese companies using a type of purposive sampling. The selection criteria required sample companies to be located in Taiwan and to compete globally to confirm that intellectual capital is positively associated with a firm’s global performance and a moderating effect of knowledge management strategy on the relationship between intellectual capital and global performance [13].

Zekić-Sušac, M., & Has, A. claimed that all previous research about integrating DM in KM has shown success of data mining methods in marketing, but the integration in a knowledge management system is still need more investigated. So, they suggested an integration of two data mining techniques: association rules and neural networks in marketing modeling to integrate with knowledge management and produce better marketing decisions[14].
Knowledge management and data mining techniques are really useful for marketing especially for organizations which have huge amount purchase transactions. Knowledge management and data mining also can help to increase the profit because of the correct decisions made by marketers. Al Essa, A., & Bach, C. showed how knowledge management and data mining can be used to provide better answers from huge amount of data of customers and purchase transactions. The goal of their article is to demonstrate the importance of using knowledge management and data mining for supporting marketing decisions. It shows how the data mining techniques and tools can extract hidden purchase patterns that can help to make better decisions by the marketers [15].

Despite of the valuable contributions of the previous related works in helping us recognize and refine our research methodology, our research, focus on the importance of integrating knowledge discovery as DM and knowledge management processes especially in domain name marketing. Also we would like to improve (.PS) domain name as an intellectual capital asset which represent the Palestinian Identity, Palestinians history, culture, lifestyle and further to the Palestine home land issue on the cyber and internet.

III METHODOLOGY

In this section, we are going to describe data mining techniques used in this research including data preparation and representation, association rules and data classification, as well as a description of how we integrated DM with KM to gain knowledge and its usage for improving marketing. Figure shows the integrated methodology which consists of two main phases, Data Mining Phase and Knowledge Management Phase. A detailed description of the methodology and its phases is discussed in section D.

A Data Preparation and Representation

The Palestinian National Internet Naming Authority (PNINA) – the Registry Organization for .PS CcTLD - uses an open source domain registry system called Cocca registry system. The Dataset Contains 12 attribute with 489 record for the 2 and 3 character and premium domains which has a high value in terms of registration fees, containing all domain registration information about the registrar and the owner of the domain including domain named, Registrar name, ID, Email, Phones, Country TLD, Address and the same information for the owner [22]. PNINA has 96 registrar Companies, 21 of them are International and the others are spread over the whole of Palestine. The Cocca registry system adds an attribute for each registrar and owner as CcTLD, which identifies the Nationality of the registrar company, and the owner of the domain according to other attributes like Phone, Fax or address as shown Figure 4.

We used a number of operators that satisfies the improvement of the data set to be ready for next mining techniques such as Replace Messing Values, Filtering, Sampling, Outlearning algorithms and Attributes selection.

B Applying Association Rules (AR’s)

We use AR’s data mining technique for discovering relations between patterns in large database to identify strong rules that could be used to capture knowledge [16]. In our case we can predict that AR may give us a good result when taking Owner’s_country_TLD attribute as a label to generate all AR’s according to the nationality of the owner.

C Applying Classification Methods

We applied k-Nearest Neighbor classification method (k-NN) data mining technique on our data set which we labeled the "Owner_country_TLD". This may help us to find which countries or regions around the world interested in registering the Palestinian CcTLD. k-Nearest Neighbors algorithm (k-NN for) is a nonparametric method used for classification and regression[17]. The input is a feature node which will be defined according of the similarity to the nearest k nodes in the trained feature space. In Classification, the output is class membership. Input node is classified by the neighbor’s majority vote then it will be assigned to the class that contains the most k nearest neighbors. k is a positive integer that has to be given and known for the input which represents the number of nodes that the input will be processed with to find its best class, accuracy of the k-NN algorithm depends on k and k may depend on the structure of the dataset itself [18, 19].
D Integrating DM and KM (DKI Model)

We propose a methodology that shows how we integrated DM techniques in KM to generate a good marketing plan to improve Premium Domain Marketing as shown in Figure.

We introduced our methodology as two interconnected phases: Data Mining (the knowledge discovery) and Knowledge Management (sharing and planning). Each phase consist of two steps beside data preparation. These four steps should be accomplished respectively.

The Data Mining Phase consists of two steps. **Step 1** is applying Association Rules (ARs) to discover Interesting patterns that should give a vision of relationship between Registrars and customers. AR’s provide patterns in the form of X → Y (if X then Y) where X is the registrar and Y is the Customer and vice versa. **Step 2** is the Classification method where the dataset is classified using an algorithm known as k-Nearest Neighbor classification method (k-NN) that will classify customers according to their nationality to find the most attention region of the world about registering premium domain names.

Knowledge Management Phase also consists of two steps. **Step 1** is about sharing interesting patterns and classification results about customer profile with employees or with a set of employees involved in knowledge management to collect new ideas, rank these ideas and the select the ideas that could be transformed into a good marketing strategies, this step should achieve visibility and reasoning to enable employees to be actively participate in generating and innovating new marketing plans that could achieve the main goal and develop the organization environment and learning. **Step 2** is usually managed by marketing and sells managers or the General Manager - as in PNINA - as well as including other professional employees. To generate real marketing plans that could achieve the overall goal for the organization which is improving registry for premium domain names, mainly focus on the following types of marketing strategies: (1) Using media and techniques for product promotion, (2) Revision of pricing such domain names, (3) Develop a full package for the domain name like offering the domain with a strategy of use with an idea of an innovation project, (4) Gather all expected names in a domain field and promote it directly to the targeted customers, (5) Find new registrars in neglected regions and (6) Offer domain names according to countries and people culture and polices. This step effects should increase the sale, cross-selling index and competitiveness of the company.

IV RESULTS AND DISCUSSION

For this section, we are going to list all results that we got from each step of our two phases in DKI Model that was described before in methodology section.

A Phase 1: Data Mining (DM)

**Step 1** Figure shows two selected association rules; we can see that Palestinian Owners cooperate with local Registrars for registering domains. Also from the second selected association rule, we conclude the all local registrars are selling these domains for Palestinian customers.

![Figure 5: Gained Association Rules](image)

For the 1st rule, we can explain it because of the price range that PNINA sells domains for Local and international registrars. People try to buy with the least price.

For the 2nd rule, we can predict that it’s due to the Weakness in marketing and technical capacity. Most of local registrars do not use a full-automated registration system that helps customers to buy and modify their domains.

**Step 2** K-NN Classifier Confusion Matrix as in Figure shows an accuracy up to 96.5%, which is a good number that we can count on it. K-NN built 10 classes of classification depending on country location.

![Figure 6: K-NN Confusion Matrix](image)

The count graph as in Figure shows that local sales have the biggest amount, US comes second and Switzerland comes third. Others come last which are China (CN), Great Britain (GB), Germany (DE), Indonesia (ID), Czech Republic (CZ), Norway (NO) and Sweden (SE).

![Figure 7: K-NN Graphic Count](image)

We can conclude that we need to concentrate about global marketing, especially in Europe and the US. Other places like Africa, Latin America and Australia can be a future work.
B Phase 2: Knowledge Management (KM)

Step1 The activities about sharing interesting patterns, classification results about customer profile with employees or with a set of employees to collect new ideas, rank these ideas and select the native ideas can be implemented by the help of KM software tools that are available today, the choice of a tool should depend on functionalities needed and its cost. Sharing is able to contribute organization environment and learning, generate some new ideas on marketing strategies and achieve the main overall goal.

Step2 Appointed Managers and officials should take this step seriously because we count on it to generate real marketing plans that could achieve the overall goal for the organization.

V Recommendations

As a result of the four steps in the two phases of the methodology and as results shown in figure 7, the registry of (.PS) domain is concentrated in Palestine. This draws the attention to the fact that Using media and techniques for product promotion should increase targeted audience worldwide and with reference to the types of marketing strategies mentioned in subsection III.D we present the following recommendations:

- Using media and techniques for product promotion such as:
  - Sponsoring contests
  - Registrar referral incentive as a way of encouragement
  - Promotion while supporting causes and charity
  - Branded promotional gifts
  - Registrar and customer appreciation events
  - Customer surveys and feedback
- Revision of pricing such domain names to increase selling and encourage customers to get it.
- Develop a full package for the domain name like offering the domain with a strategy of use with an idea of an innovation project.
- Gather all expected names in a domain field and promote it directly to the targeted customers.
- Find new registrars in neglected regions and thus create new points of sale to facilitate dealing with customer
- Offer domain names according to countries and people culture and polices.
- Identifying renowned brands in each market area and begin a promotion plan for its names under the CcTLD (.PS)

VI Conclusion

Integration Data Mining techniques such as Association Rules and Classification methods in Knowledge Management leads to facilitate the process of discovering knowledge and decision making for generating marketing strategies and solutions. This approach is used in a previous research. The originality here is applying the technique to the (.PS) domain name as a case study. As result of low prices of local registers, we can conclude that, local registrars register more than 94% of the (.PS) domains; the variation of prices between international and local registrars gives a big chance for locals to sell more domains. However, international customers do not register from a local, which is because of the Weakness in marketing and technical capacity. Local registrars do not use a full-automated registration system that helps customers to buy and modify their domains. The classification methodology shows that the most attention region for registering such domains is represented in Palestine, small ratio can be noticeable in the US and Switzerland (CH). Also we can find some registered in China (CN), Great Britain (GB), Germany (DE), Indonesia (ID), Czech Republic (CZ), Norway (NO) and Sweden (SE). This means that we still need to find good marketing plans for such domains in Europe, Asia and the USA. Some markets do not exist like Africa, Australia, Canada and South America. These markets must be exploited well in order to expand the target customers and thus increase sales.

References

Ibrahim S. Alfayoumi, Wael W. Alsarraj / Improving Premium Domain Names Registration for .PS Domain CcTLD Basing on Knowledge Discovery and Data Mining Techniques (2017)


