Targeted Online Marketing using Social Networking

Article *in* International Journal of Online Marketing · September 2015

DOI: 10.4018/ijom.2013070103

3 authors, including:

Mohamed K. Watfa

80 PUBLICATIONS **926** CITATIONS

SEE PROFILE

Mahmoud Bakkar

Abu Dhabi University

20 PUBLICATIONS **21** CITATIONS

SEE PROFILE

Targeted Online Marketing using Social Networking

Mohamed K. Watfa, Faculty of Engineering & Information Sciences, University of Wollongong in Dubai, Dubai, UAE

Nima Najafi, Faculty of Engineering & Information Sciences, University of Wollongong in Dubai, Dubai, UAE

Mahmoud Bakkar, Faculty of Engineering & Information Sciences, University of Wollongong in Dubai, Dubai, UAE

ABSTRACT

For the past century, companies have had the luxury of deciding what they will produce and sell, what their brand message will be and how they will deliver it to their audience. Planning how best to allocate marketing dollars is arguably the most annoying challenge marketers face. The social web and more specifically, Web 2.0 have changed the original marketing strategy and have done so by giving rise to a new way of marketing where people belong to different social groups and markets have become conversations or recommendations. Social networking sites such as Facebook have reported exponential growth rates and have attracted millions of registered users, and they are interesting from a marketing point of view because they store large amounts of sensitive personal user data. In this paper, the authors introduce a targeted Marketing strategy that exploits group membership information that is available on social networking sites. More precisely, the authors show that information about the group memberships of a user can lead to a more efficient target marketing campaign. To determine the group membership of a user, the authors leverage well-known web browser history recording attacks and other available crawling services. The authors' proposed algorithm is designed to use the captured customers' details and generate target marketing campaigns by relating each customer with certain rank of products. The authors demonstrate an Experiment by sending a SPAM Email to more than one thousand Facebook users and relate them with certain product pages. To measure the efficiency of the proposed techniques, the authors analyzed how many people have consequently accessed those pages and clicked on the products links inside those pages to get a promising success factor of 82%.

Keywords: *Facebook, Social Networking, Targeted Marketing, Web 2.0, Web Browser History*

1. INTRODUCTION

Marketing research has always looked for specific components of proposed marketing plans that are performing better than others. These components can be the missing recipe which can forge a stronger bond between marketers and consumers, thus increasing the consistency, speed and control of brand communications leading to a competitive advantage over other organizations. Unfortunately, the vast majority of marketers lack the time, data, wherewithal,

DOI: 10.4018/ijom.2013070103

or inclination to create this knowledge. As a result, marketers are forced to allocate marketing investments in a vacuum using outdated strategies and tactics such as segmentation of markets and then targeting all customers in a particular region or people with a particular interest. Recent changes in media have made it both undesirable, and effectively impossible, to reach a given target audience with conventional mass-marketing tactics.

Many different flavors of online communities have surfaced over the years. However within the last few years, newly created communities are offering more rich interaction. These marketing strategies allow conversation, connection, and a sense of community among its members. If marketing content can be sent to those interested, that will lead to steadiness, swiftness and direct control over brand communications. The consumers who remain virtually connected to the advertised company's link are the ones who are most interested in who the company is, and what it has to offer and therefore more focused marketing of these groups are preferred.

This paper proposes techniques to market the right product to the most suitable group of online users using the data gathered from their group memberships and their corresponding social graph. Traditional offline channels of marketing products and services are important functions of business activities and more importantly extending these channels to include electronic marketing or e-marketing activities requires appropriate preparation, planning, and careful implementation. There are a number of different ways a marketer can brand their products or services. Yet, if the right advertisement reaches the right potential customer through the right channel at the right time, then competitive advantage and boosted sales will be a definite result. Marketers are always on the lookout for new data about the market. An excellent platform for Marketing is the World Wide Web in general and Social Networking platforms in specific. Social Networking can help the marketer identify a targeted marketing campaign based on almost true information gathered from people's profile which are publicly accessible.

The current marketing strategies are the traditional ones which are used regularly and have been used over the years. First, markets are segmented - which is that markets are broken into fragments or parts by different aspects of customers like demographics, geographies, customer tastes and preferences. After segmenting, some of the market segments are chosen which have some potential for a company which is going to market a product or service to the customers. This is called target marketing, where markets are targeted. Once the target markets are chosen, these markets are positioned relatively with the products or services by a company in relation to competitors products and this way a product or a service occupies the mind of a customer and the marketing campaign reaches the customer which was part of the targeted market. Now with Web 2.0 and Web 3.0, new and effective marketing techniques are coming up and are being discussed. In this paper we will also discuss a new marketing technique which is marketing through social networks like FaceBook. In this paper, we discuss a set of Algorithms to extract customer's details from social networking websites such as FaceBook and then store it in a database which has the design similar to those in companies like Google. The data in these tables are based on the users' interest and the weight of their usage. Based on the stored data, we will try to intersect the customers details to customize the online campaigns and identify the key users for certain products. After the right target is well identified, a targeted spam is initiated which has much less network consumption than traditional spam marketing strategies.

The rest of this paper is organized as follows. Section 2 introduces the inter-browsing history recording techniques. Section 3 discussed some lessons learnt from Google and how they can be utilized for social networking targeted marketing. Section 4 presents our proposed algorithms and Section 5 shows some experimental results. Section 6 presents

some research work related to the proposed techniques. We conclude this paper in Section 7 with some possible future directions.

2. INTERNET BROWSERS HISTORY RECORDING TECHNIQUE

2.1. De-Anonymization Attack Overview

The technique proposed in this paper tries to de-anonymize the information about the anonymous users of social networking websites by taking advantage of different techniques such as recording the users internet browsing history through some crawling services in order to advertise the most appropriate products or services to those users who might be interested. Let us first define some terminology that will be the foundation of our proposed marketing strategies:

- **Definition 1.1 - Social Networks:** A social network *S* is defined as a graph with a set of vertices V and a set of edges E. In this graph S, the vertices of S, V(S), represents the social networking users, and the edges in S, E(S), represents the "friendship" relation between two users;
- **Definition 1.2 - Groups:** The social network contains a number of groups with a number of users subscribed. Let G represent the set of all the groups in the social network. Each group $g \in G$ contains a set of users the social network i.e. $\forall g \in G$: $V(g) \subseteq V(S)$. Social networks typically do not allow empty groups without any users (and social network actively deletes such groups). Thus, we can assume, without loss that: $\forall g \in G$: $g \neq \varnothing$;
- **Definition 1.3 - Group Membership:** Each user $v \in V$ is a member of *n* groups, with $n \geq 0$. We can therefore build a two dimensional matrix *GM[a][b]* representing the group membership of *a* users in *b* groups in a social network:

$$GM[i][j] = \begin{cases} 0 & if\ user\ i\ does\ not\ belong\ to\ Group\ j \\ 1 & if\ user\ i\ belong\ to\ Group\ j \end{cases}$$

Note that *n* would therefore represent the number of 1's in a specific row and the $n < a$. Also, if $n = 0$ (i.e., the user is not a member in any group), the row contains only zeros. This is the worst case for our marketing campaign as not keywords can be linked to specific users. The matrix *GM* will be utilized to fingerprint specific user interests and market accordingly (Table 1).

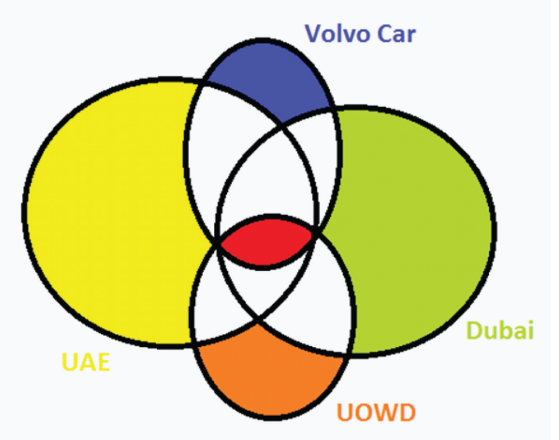
2.2. Grouping and Target Marketing

In order to have a more targeted marketing approach, findings about the group membership of a social network users can be very helpful. To make the scenario clear, an example might be helpful. Imagine, Mike is a member of "We love UAE and Dubai" group. He is also a member of UOWD university, and Volvo car fans group. Figure 1 shows how we can target

Table 1. Matrix GM

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
User 1	0	0	0	1	0	1	1
User 2	1	0	0	0	0	0	0
User 3	1	1	1	1	1	1	1
User 4	0	0	0	0	0	0	0
User 5	1	0	1	0	1	0	1

Figure 1. Group intersection for target marketing. The red zone represents the overlap between several group memberships and indicates the set for which the user belongs to.



such a person and narrow down the targeted person with much less search effort.

In this context, we can utilize the GM matrix to narrow down a specific identity. For example, User 7 is the only member belonging to all seven groups and therefore can be easily identified.

2.3. Browsing History Recording Technique

This technique is very dependent on the browsers which the users use. This way we will access the user's browsing history and find out their memberships in groups on social networking websites. The worst case scenario in this technique is when a user has a good amount of technical knowledge and is conscious about his privacy and security and he clears his browsing history manually. Different internet browsers delete the browsing history after a specific period. By default, Internet Explorer deletes its history after 20 days, Mozilla Firefox after 90 days, Apple Safari after 30 days, while Google Chrome never deletes the browsing history of the user. We understand that Google also wants to use the browsing history of users for targeted advertising and that is why they designed a web browser that never deletes the history, especially the cookies. As researched previously, many techniques are available for capturing the browsing history. The overall idea is as follows. An attacker creates an HTML page with links to the targeted web pages of interest using background image tags in the a: visited style information. The image can be referenced with URLs and the user's browser will then access these URLs if a target web page is in the browsing history. It can also run a script on the user's side and send all the browsing history of the target to the attacker as is shown in Figure 2.

The result that the attacker will receive will be similar to the one shown in Figure 3.

At this stage, the information needed for a dedicated target market is there available for the attacker to view and make the next step to segment the target and start spamming. The most important and essential return for the attacker is the dynamic token which is the groupID in this scenario. Even if the website is using SSL (secure socket layer) and since the attacker only needs the client side to be hacked, the groupID can be sent. By finding the groupID of a social network user, the attacker almost knows the interests of the victim, and by having so many different interests of one victim, the attacker can target him in the most possible niche way.

Figure 2. Browsing history attack scenario

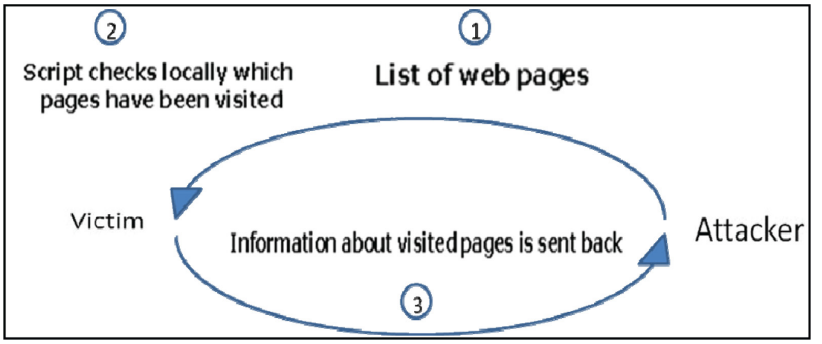


Figure 3. Result samples received by attacker

1.http://www.facebook.com/home.php?ref=home+
2.http://www.facebook.com/group.php?gid=[groupID]&v=info&ref=nf+
3.http://www.facebook.com/ajax/profile/picture/upload.php?=[userID]+
4.https://www.xing.com/net/[groupID]/forums+

But in the case of a user deleting their browsing history or if the history is not accessible, we use Crawling Service which is discussed in the next section.

2.4. Crawling Service

A web crawling software is addressed in Mozenda (2012) and is described as a piece of computer program usually called a web crawler or a web spider. These types of programs search the websites available or addressed on the internet and make copies of their source code. Once the copies are made, another set of applications can be used to separate and index the data on the targeted web sites. This software can be set to search the internet using specific parameters to help research and data analysis and sometimes can be used by an attacker to find relevant information for possible spamming. These programs such as "Mozenda" or "80legs" also find contact information such as addresses, phone numbers, business names, and email only if they are allowed to be shown publicly. Based on Legs Inc (2011), the attacker cannot access that information that is set to private in social networking websites. Once the information is found and copied, it can then be manipulated to suit the needs of the attacker. These services are available and apparently not very expensive to use. Suppliers claim that a couple of hundred dollars that are spent per year to access this information is actually worth millions of dollars. Either the information the attacker needs is gathered by the history recording technique or by web crawling services, he can compile a database of victims which will be shown in next chapter of this paper.

3. LESSONS FROM GOOGLE

The best tried and tested experience is done by Google in terms of targeted advertising. In this part of the paper we try to imitate what they have done in terms of the algorithm and database design and match it with our data sets. According to (Google Inc., 2011) Google

automatically delivers targeted advertisements to the content or audience. Google does it using the following several ways:

- **Contextual targeting:** A technology that uses such factors as keyword analysis, word frequency, font size, and the link structure of a webpage, in order to find out what that page is about and precisely match the Google ads to each page;
- **Placement targeting:** With placement targeting, advertisers choose specific ad placements, or subsections of publisher websites, on which to run their ads. Ads that are placement-targeted may not be precisely related to the content of a page, but are hand-picked by advertisers who've determined a match between what your users are interested in and what they have to offer;
- **Interest-based advertising:** Interest-based advertising enables advertisers to reach users based on their interests (e.g. 'sports enthusiasts'), and allows them to show ads based on a user's previous interactions with them, such as visits to advertiser websites.

Google uses the above three techniques to target users. Apparently the most relevant technique to this paper's case is the last one which is Internet-based advertising. Obviously, Google recognizes users' interests based on previous actions that the specific user had taken before. Here, we want to focus on technical terms of how Google does internet based advertising.

3.1. Google Techniques

Previously known as the Google Content Network, the Google Display Network includes a collection of websites that have partnered with Google or Google properties such as YouTube that display AdWords ads. These ads, which can be in any format, can be targeted automatically. Websites that join the Google Display Network allow Google to put ads on their websites and also benefit to be a part of the list that Google will recommend users to access to. Based on (Google Inc., 2011), Google uses cookies to keep users information in order to target the right people for internet-based advertising. The cookies are saved in users internet browsers and it can be one of the reasons that the Google Chrome browser never deletes the browsing history, even though it makes the browser less secure. Google categorizes the interests into: Arts & Entertainment, Autos, Beauty, Books, Business, Electronics, Finance, Food & Drink, Games, Hobbies & Leisure, Home & Garden, Internet & Telecom, Jobs & Education, Law & Government, News, Online Communities, People & Society, Pets & Animals, Real Estate, Reference, Science, Shopping, Sports Add, Travel Add, and World Localities Adds. Our proposed technique does not use cookies as Google does but rather uses both a crawling service and group membership information to target a specific group of users. Using similar interest categories, we will design our database.

3.2. Database Design

We designed our database to have five tables *Customers, Product Category, Interest keyword frequency, Keyword type* and *product Rank Frequency* (Figure 4). As we can see the relationship between the customers and the product Category is many-to-many and every customer has an interest keyword frequency and every keyword has many product categories. To identify our keywords we decide to master the category types into three main types that represent everything in our life *"LAND, SKY, SEA"*. For example, the customer searches for an Airline Ticket so when we record his keywords in our database, we have the keyword Ticket but the customer was always purchasing cinema Tickets online so that Ticket could be Airline or it could be Cinema and to decide that we designed the table *Keyword type* so if the keyword is categorized as LAND it means cinema and if its SKY it means Airline Ticket and if we have two Tickets linked with the LAND type we use the third

Figure 4. (a) The database Relationship; (b) Snapshot of Customer database table; (c) Snapshot of the Intersect Keyword Frequency; (d) Snapshot of the Product Category database table; (e) Snapshot of the Product Rank Frequency database table

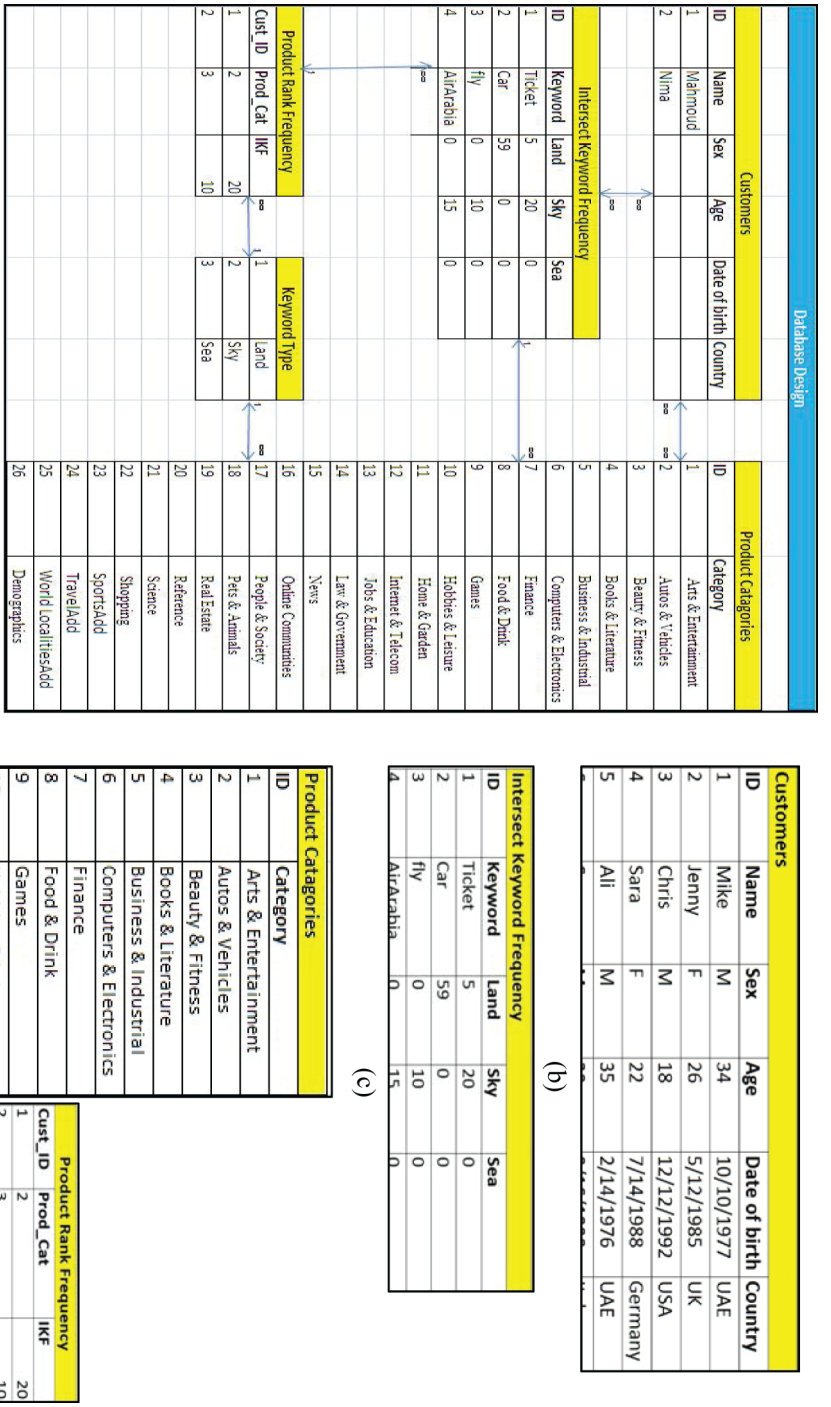


table which is *product rank frequency* to decide which product category the keyword belongs to. We use that table as a dataset to act as the input keywords for the algorithm.

4. ALGORITHM AND IMPLEMENTATION

After we gather the customers interests in our database using the history recording technique or crawling services, we now discuss using the *Evolutionary Local Selection Algorithm (ELSA)* (Kim, Street, & Menczer, 2001) to optimize our searching of the targeted keywords designed to intersect the customers details and generate the targeted marketing campaigns and relate each customer with certain rank of products. These products are ranked according to the customers' interest using *ELSA algorithm*. This will dramatically increase the efficiency of e-marketing and reduce the network traffic by customizing the online campaigns according to the highest rank in the customers products rank frequency table in our database and relate those customers with these campaigns accordingly.

The procedure of ELSA expressed in pseudo code is given in Figure 5.

The ELSA algorithm (Figure 6) will be used to optimize our searched targeted customer to a small subset of the big set of customers using the keywords that were matched with certain features (for example: Volvo RED Color). This will therefore find the subset of customers interested in buying a Volvo car in red color and as a result will link this subset of customers with our online campaign to sell them a red Volvo car. In our experiment, we use the second product category Auto & Vehicles from the product category in our Keywords Database to identify the customers interested in buying

Figure 5. The process of integrating the social network users to identify the optimal keywords subset using ELSA algorithm

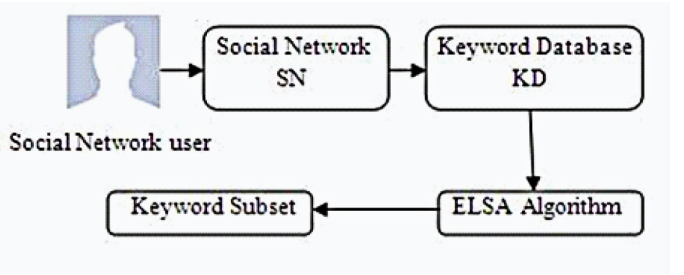


Figure 6. A modified procedure using Evolutionary Local Selection Algorithm (ELSA)

```
Initialize population
Apply local search
if a solution is found then return "a satisfying assignment is found"
while (stopping criteria is not reached & no satisfying assignment found) do
              Select parents
              Recombine pairs of parents
              Apply local search
              if the offspring is determined as a local maximum point then
                            Mutate offspring
                            Apply local search
              end if
Replace parents by offspring
end do
return "no satisfying assignment found"
```

Auto then we use the *ELSA Algorithm* to give us the small subset of customers interested in RED Volvo cars.

So, the ELSA Algorithm will return the Top *x%* of customers among the Auto & Vehicles group who are most likely to buy a red Volvo car. We use the Artificial Neural Network to extract the predictive information from each subset. In Figure 7 we input the whole set of Keywords that are related to all customers and we input our features search "Volvo, Car, Red" to identify the small subset of customers who are interested in buying a Volvo Car by using the ELSA Algorithm then we use the Artificial Neural Network Algorithm to predict the customers who are, for example staying in Dubai and give us the estimated subset of customers who are living in Dubai and are interested in a red Volvo car.

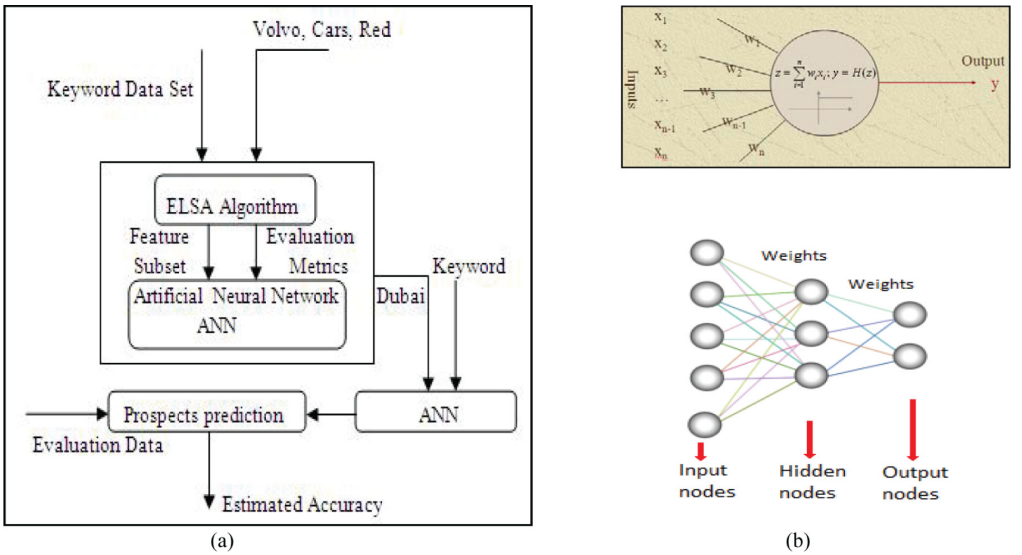
The *Artificial Neural Network (ANN)* is an algorithm derived from the nervous system in the human body to act in the same behavior as the nervous system and it is used for the final prediction of the products that is related with the customer's highly frequent keywords and then specify the product category and target it in our online campaign and the main purpose of the artificial neural network is to transform the input into a meaningful output. Using a similar model depicted in Figure 7 (b), we use *McCullogh-Pitts neuron* and represent *X* as the keyword that is stored in our database and *W* as the keyword frequency from different product categories.

A nonlinear generalization of the *McCullogh-Pitts* neuron is given by $y = f(x, w)$ where y is the neuron's output, x is the vector of inputs, and w is the vector of synaptic weights. To ease computations, we utilize a unit step function model where the output is set at one of two levels, depending on whether the total input is greater than or less than some threshold value (Figure 8).

5. EXPERIMENTAL RESULTS AND ANALYSIS

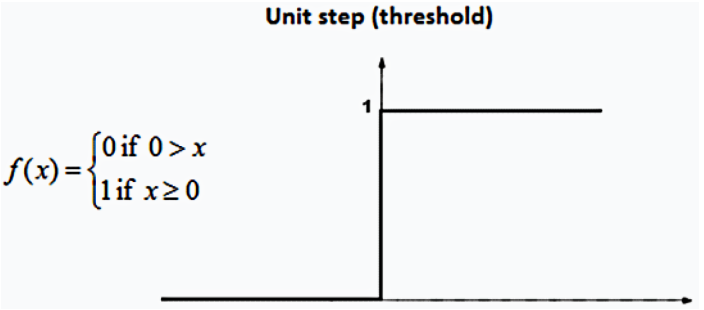
We demonstrate an experiment by sending a Spam Message to 1000 FaceBook users and relate them with certain product pages that were customized based on our database of the

Figure 7. (a) Estimated Accuracy of the Algorithm; (b) The McCullogh-Pitts model and an Artificial Neural Network (ANN)



(a) (b)

Figure 8. A unit step function for the McCullogh-Pitts neuron



keywords. We were able to successfully record their history and build their keywords records in our database. We also successfully captured their group memberships and analyzed the overlap. We analyze how many people accessed those pages and how many people clicked on the product links inside the pages and which product links were clicked. We also analyze the number of clicks inside each page and rank those links accordingly with certain products. The following page/product hit results were captured as depicted in Figures 9 and 10.

To further analyze the performance of our proposed algorithms, we also measure the success of the targeted 1000 messages which were carefully selected based on related keywords that we have taken from their captured group membership overlap as discussed in previous sections. Every person who actually visited the page we had sent them within a specified period of time is considered a success. Figure 11, depicts the utilization of our proposed techniques within a specific month (4 weeks).

A number of performance measurements that are evident from the experimental analysis include:

1. **Average Response Time:** 21 days;
2. **Worse Case Response Time:** 28 days;
3. **Percentage Response:** The performance of our targeted marketing approach is:

$$\frac{\left\{\left(\frac{120}{150}\right)+\left(\frac{200}{250}\right)+\left(\frac{310}{350}\right)+\left(\frac{450}{550}\right)\right\}}{4} = 82.5\%$$

Using the proposed techniques, the performance results can be considered as a significant improvement over traditional techniques and can be considered very encouraging in the

Figure 9. The results of sending 1000 spam email link

Experiment Results		
Time	Pages Hits	Products link Hits
1 week	150	120
2 week	250	200
3 week	350	310
4 Week	550	450

Figure 10. Display the pages hit and product link hits

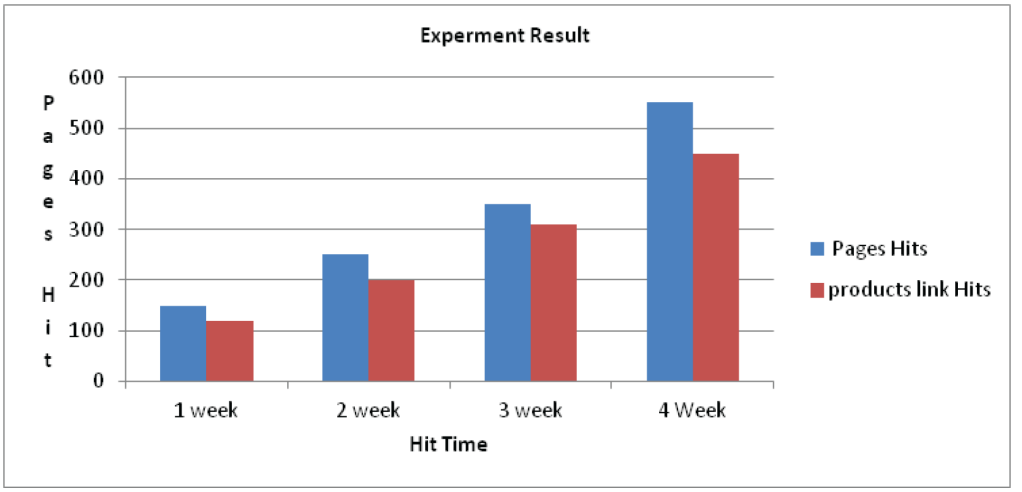
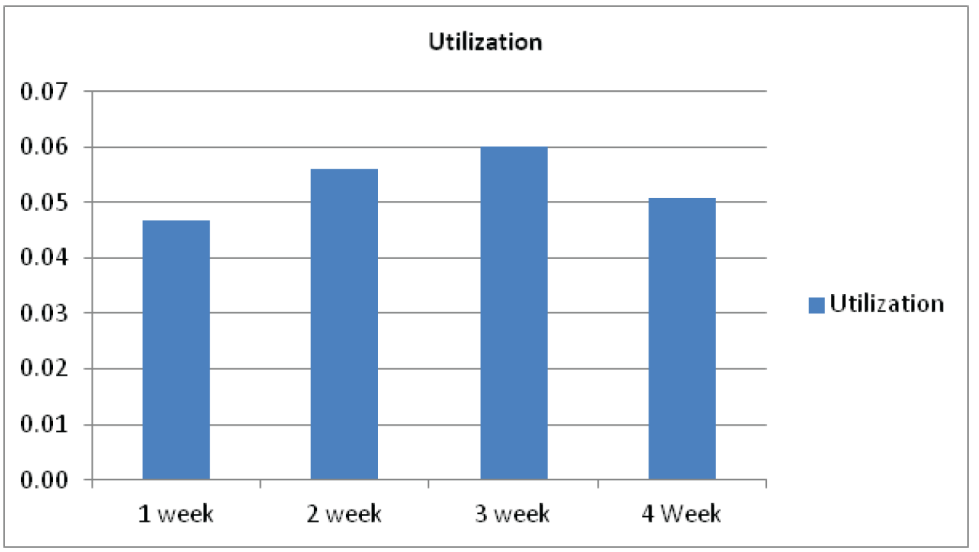


Figure 11. The utilization of our web pages over four consecutive weeks



areas of marketing. Also, the quick response from some users is clearly evident using our technique. Some other formats of less targeted marketing approaches such as television advertising and billboard advertising usually show a less than 5% overall response and success, which would therefore imply the evident success of our proposed techniques.

6. RELATED WORKS

Social networking websites such as FaceBook have done a great job connecting people to each other and at the same time claiming to have hundreds of millions of those users having all the needed data for advertising rather than any other third party. FaceBook's approach to ad-

vertise on its website is different than Google's approach and ours. The model is useful for individuals and businesses that have a profile on Facebook. The service is called "Facebook Advert". Some research indicate that it is not suitable for older generations. Yet the Facebook team claim that interestingly, while Facebook started on a university campus (among a young generation of students), the fastest growing FaceBook users worldwide is among people 35 years and older. Dan Hobin, CEO of G5 Search Marketing says: "FaceBook helped our client target a specific audience and offer a solution for their specific needs. Offering college students self storage right at the end of the semester with Facebook Adverts helped StorQuest post record-breaking student occupancy rates." Based on what is shown on FaceBook, the results from StorQuest's Facebook advert campaign were outstanding.

A very useful aspect of FaceBook Advert (unlike the Google techniques) is that the advertisers can be mostly sure that the targeted person is real and exists. Also, in (Kimura, M, Saito K., and Nakano R., 2007), the authors also adopted another strategy to extract marketing oriented useful nodes in a network. They proposed a method of estimating marginal gains on the basis of a graph theory. They found that if they build up the social network with more related information about customer's favorites and feedbacks then the basic structure of a social network such as degree of centrality can be a good method for identifying the most influential users in an online community. According to Zang, Wang, and Xia (2010), the researchers believe that the best marketing strategy is to let influential customers be the marketers. Therefore, they concentrated on identifying influential reviewers that might affect a large group of their friends, connections and other online customers to buy the product. They also tried to come up with algorithms and actual implementations.

Several publications have analyzed and measured features of social networks that are privacy-related. Griffith et al. (2005) use public records to infer individuals' mothers' maiden names, and Heatherly et al. (2013), as well as Zheleva et al. (2009), show how public data provided by social networks can be used to infer private information. Mislove et al. present a measurement study on social networks (Mislove, Marcon, Gummadi, Druschel, & Bhattacharjee, 2007) while Bonneau and Preibusch evaluate the privacy settings and policies of a large number of social networks in Bonneau and Preibusch (2009). Closely related to this context, several recent papers focus on scenarios for malicious activity directed against social networks (Brown, Howe, Ihbe, Prakash, & Borders K., 2008). Bilge et al. (2009) shows the feasibility of automated identity theft attacks in social networks.

The de-anonymization scenario presented in this work leverages a browsing history stealing technique that is based on CSS and has been known since the year 2000. This technique has been discussed in several browser bug reports (Ruderman, 2000) and has been shown to be practical for targeted phishing attacks in (Jakobsson & Stamm, 2006). Despite its malicious potential, browser history stealing has not lead to any changes in browser software. Felten and Schneider show an attack on web browsing history by analyzing caching operations (Felten & Schneider, 2000), while Bortz et al. (2007) use timing attacks to recover private information from web applications. I was also shown that statistical methods can be applied to de-anonymize micro-data by cross-correlating multiple datasets and that it is possible to de-anonymize members by mapping known, auxiliary information on the social network topology (Narayanan A. and Shmatikov V., 2009). In Diaz, Troncoso, and Serjantov (2008), Diaz et al. present a de-anonymization approach that uses information gained from observing communication patterns between social network members.

7. CONCLUSION AND FUTURE DIRECTIONS

Marketing wisdom says, if a product is marketed at the right time, to the right target through the

right channel, success will be the definite result. Internet advertising - and to be very specific, social networking websites can provide a very reliable infrastructure for marketers to find the right people as their target and segment markets in one hand and advertise to them through the same channel or even other channels based on the users interest. In this paper, several techniques to gather useful information from a social network were rigorously explained. Our proposed techniques combine web browsing history to get group memberships which in turn lead to the design of an efficient database of users using Artificial Neural Networks. The database created can therefore be used to market specific products liked to specific keywords. Our results are very promising compared to traditional marketing methodologies as our proposed techniques showed a success percentage of 82%.

Although our work shows great promise for the future of marketing but it still lacks a number of critical realistic measures. These include the ethical and privacy issues behind the use of web crawlers that would typically be running without users consent. Also, the algorithms can be further optimized by compiling more information about specific users such as IP Addresses, location, gender and other that might give better targeted marketing results. Crawling data in social networks is an ethically sensitive area. Clearly, one question that arises is if it is ethically acceptable and justifiable to conduct crawling experiments in social networks. This is a big research area that needs further discussions.

REFERENCES

Bilge, L., Strufe, T., Balzarotti, D., & Kirda, E. (2009). All your contacts are belong to us: Automated identity theft attacks on social networks. In *Proceedings of the 18th International Conference on World Wide Web (WWW)* (pp. 551-560).

Bonneau, J., & Preibusch, S. (2009). The privacy jungle: On the market for privacy in social networks. In *Proceedings of the Eighth Workshop on the Economics of Information Security (WEIS),* (pp. 1-6).

Bortz, A., & Boneh, D. (2007). Exposing private information by timing web applications. In *Proceedings of the 16th International Conference on World Wide Web (WWW)* (pp. 252-259).

Brown, G., Howe, T., Ihbe, M., Prakash, A., & Borders, K. (2008). Social networks and context-aware spam. In *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work (CSCW)* (pp. 403-412).

Diaz, C., Troncoso, C., & Serjantov, A. (2008). On the impact of social network profiling on anonymity. In *Proceedings of the 8th International Symposium on Privacy Enhancing Technologies (PETS)* (pp. 44-62).

Facebook. (n.d.). Retrieved July, 2011, from http://www.facebook.com/adsmarketing/index.php?sk=targeting

Felten, E. W., & Schneider, M. (2000). Timing attacks on web privacy. In *Proceedings of the 7th ACM Conference on Computer and Communications Security (CCS)* (pp. 25-32).

Google Inc. (2011). Retrieved July 2011, from https://www.google.com/adsense/support/bin/answer.py?hl=en&answer=9713

Google Inc. (2011). Retrieved July, 2011, from http://www.google.com/ads/preferences/html/about.html

Griffith, V., & Jakobsson, M. (2005). Messing with Texas, deriving mother's maiden names using public records. In *Proceedings of the Third Conference on Applied Cryptography and Network Security (ACNS)* (pp. 91-103).

Heatherly, R., Kantarcioglu, M., & Thuraisin, B. (2013). Preventing private information inference attacks on social networks. *IEEE Transactions on Knowledge and Data Engineering*, *25*(8), 1849–1862. doi:10.1109/TKDE.2012.120

Jakobsson, M., & Stamm, S. (2006). Invasive browser sniffing and countermeasures. In *Proceedings of the 15th International World Wide Web Conference* (pp. 523-532).

Kim, Y., Street, N., & Menczer, F. (2001). *An evolutionary multi-objective local selection algorithm for customer targeting* (pp. 759–766). IEEE.

Kimura, M., Saito, K., & Nakano, R. (2007). Extracting influential nodes for information diffusion on a social network. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence* (pp. 1371–1376).

Legs Inc. (2011). Retrieved July 2011, from http://www.80legs.com/crawl-packages-social-networks.html

Mislove, A., Marcon, M., Gummadi, K., Druschel, P., & Bhattacharjee, B. (2007). Measurement and analysis of online social networks. In the *Proceedings of the 7th ACM SIGCOMM Internet Measurement Conference (IMC)*.

Mozenda. (2012). Retrieved July 2012, from http://www.mozenda.com/web-crawling

Narayanan, A., & Shmatikov, V. (2009). De-anonymizing social networks. In *Proceedings of the IEEE Symposium on Security and Privacy*. University of Texas at Dallas, Tech. Rep. UTDCS- 03-09.

Ruderman, J. (2000). *CSS on a: Visited can load an image and/or reveal if visitor been to a site.* Retrieved from https://bugzilla.mozilla.org/ show bug.cgi?id=57351

Zang, Y., Wang, Z., & Xia, C. (2010). Identifying key users for targeted marketing by mining online social network. In *Proceedings of the IEEE 24th International Conference on Advanced Information Networking and Applications Workshops* (pp. 644-649).

Zheleva, E., & Getoor, L. (2009). To join or not to join: The illusion of privacy in social networks with mixed public and private user profiles. In *Proceedings of the 18th International Conference on World Wide Web (WWW)*.

Mohamed K. Watfa is currently an Associate Professor in the Faculty of Engineering at the University of Wollongong in Dubai. Before that he was at the Computer Science department at the American University of Beirut (AUB). He received his Ph.D. from the School of Electrical and Computer Engineering at the University of Oklahoma in Norman, OK, USA in 2006. He was one of the youngest PhD holders to graduate from his university. He was also on the dean's honors list and was given a number of prestigious awards and scholarships. His research interests include wireless sensor networks, intelligent systems, Vehicular Ad-hoc Networks, wireless networking, resource management, energy issues, tracking, routing, and performance measures. He is the author of a number of books, the guest editor of a number of international journals, and the organizer of a number of international conferences. For more info about Prof. Watfa, check out his personal website at: http://mohamedwatfa.synthasite.com/

Nima Najafi is a software developer received a Masters in Information Technology Management from the university of Wollongong in Dubai and have extensive experience in the IT industry. He worked on research projects under the supervision of Prof. M. Watfa at the University of Wollongong in Dubai. His research interests include Software development, Network Security and Social Networking. Mahmoud Bakkar is an IT professional and received a Masters in Information Technology Management degree from the university of Wollongong in Dubai and have extensive experience in the IT industry. He worked on research projects under the supervision of Prof. M. Watfa at the University of Wollongong in Dubai. His research interests include Network Security and Privacy.

Mahmoud Bakkar is an IT professional and received a Masters in Information Technology Management degree from the university of Wollongong in Dubai and have extensive experience in the IT industry. He worked on research projects under the supervision of Prof. M. Watfa at the University of Wollongong in Dubai. His research interests include Network Security and Privacy.