

SPAM FILTERING FOR OPTIMIZATION IN INTERNET PROMOTIONS USING BAYESIAN ANALYSIS

Ion SMEUREANU¹

PhD Univ. Professor
Academy of Economic Studies
Bucharest, Pta. Romana, no. 2-5, district 1, Romania

E-mail: smeurean@ase.ro



Madalina ZURINI²

Academy of Economic Studies
Bucharest, Pta. Romana, no. 2-5, district 1, Romania

E-mail: madalina.zurini@gmail.com



Abstract: The main characteristics of an e-business and its promoting are presented. It contains ways of promoting an e-business, examined in depth the e-mail marketing principle along with advantages and disadvantages of the implementation. E-mail marketing metrics are defined for analyzing the impact on customers. A model for optimization the promoting process via email is created for reaching the threshold of profitability for electronic business. The model implements Bayesian spam filtering and applies an internal classification, using the principles of conditioned probabilities.

Key words: Bayesian spam filters; e-business; email marketing; knowledge society; metrics; Bayesian conditional probabilities

1. E-business in knowledge-based society

The knowledge society represents a new stage of human evolution, a superior quality lifestyle that involves intensive use of IT in all spheres of human activity, with major social and economic changes. Democracy, communication, understanding and cooperation are the main characteristics of this society, which makes knowledge society to be based on the multitude resources offered by Internet access.

The business environment has changed with the evolution of Internet. A web page with contact details and the activity field have almost every small firm. As a business, for most of the times, represents keeping contact with other persons, knowledge society based on the Internet brings advantages such as: presence, communication, updated information, well served clients, interfering with new techniques of propagation such as: images, sound and video.

The term of e-business is used for the first time by IBM, being defined as a "secure, flexible and integrated access for running various businesses with the combining of processes and systems that execute the basic operations of the business with those that

makes possible finding information over the Internet". E-business is more a strategy in terms of business, the main component being electronic commerce, „e-commerce“.

However, there is little information available on the Internet to help small businesses to explore the online space. Most studies simplify the structure of markets that practice e-economy by treating e-commerce companies as a generic group. The economic market encountered on the Internet can be divided into 3 structures: portal sites, market players and products and services distributors, each of which can be of 2 types: B2C, business oriented customers, or B2B, business oriented business.

Within this informatics and economic framework, the information is the one that makes the difference, and promoting an e-business is the key that opens the connections that any company needs to survive in the market.

In promoting products and services or for promotions, there are two approaches that are used in practice: mass marketing and direct marketing. Mass marketing involves media that disseminate information regarding the products through TV, radio, shops or newspapers. This type of marketing has a target group consisting in a variety of clients, there is no discrimination between them and the information is uniform transmitted. Direct marketing, on the other hand, is different from the mass one because the individual is the target group. Direct Marketing Association's (DMA) defines this type of marketing as "communication where information are systematically used to achieve quantitative objectives and where direct contact between companies and existing customers or potential ones is realized". From this definition it can be concluded that direct marketing classifies customers so that personalized advertisements and promotional activities can be targeted to a particular class of customers.

Over the past years, direct marketing increased in importance. Using the Internet has reduced operational costs of marketing and, even under conditions of low response from customers, it is sufficient to ensure success of the advertising campaigns.

Internet technologies offer more opportunities for online advertising. Most ads are generated by banners and sponsored content on web pages. However, in terms of online marketing, for a promotional campaign, email remains the most effective way of online promoting virtue of the precise destination, immediate response and incredibly low cost. Beyond this, it enables private communication and, properly used, helps building a long-term customers' confidence.

2. Differences between email marketing and spam

For an effective e-marketing strategy, all actions must be linked to the business's objectives. The main purpose of such a strategy is to have long term benefits of a competition, based on performance. This can be achieved through research, planning, experience, analysis and news in the online industry.

Site to the principles of offline marketing that are the 4P: product, price, promotion and placement, email marketing has the main characteristics: personalization, privacy, better customers' service, community, digital content, promotion sales, and security.

Promoting business means an expanded commercial see, public relationships, so that it brings the activity major advantages. In e-business, promotion is done with methods like:

- *Domain name* – as the company's web address is a suggestive name, the more quickly persons can retain it; visit more often, one that can be easily remembered. Most areas from .com or .co.uk have been sold already, so it takes ingenuity to be able to match the name, the business field of the firm with the web page's name. To increase the degree of familiarity, firms tend to find new names, not yet purchased, or use areas with other suffixes as .bizz.

- *Search engines* – Another way to encourage people to access the website of an electronic business is done through a registration on various search engines like Google, Lycos, HotBot or Yahoo. There are search engines that automatically index these websites, with links from other pages to this address. The higher the number of links, the better is the

position returned by the search engine, allowing people that seek information in a particular area to display the Internet pages most often used and spread. If you don't want this to be done manually, there are specialized firms that promote web addresses using search engines.

- *Sites for price comparison* – Web sites such as Kelkoo can be an effective way to promote an e-business so that, for every customer directed to the business, a fee is charged for the promotion site.

- *Electronic markets* – Sites such as Amazon, with external businesses presented, with stocks and prices, where you can buy directly, are useful to any small or large businesses for increasing the coverage of clients. A commission is charged for each sale through these markets, commission that does not diminish the e-business profit.

- *Banners, pop-ups and sponsored links* – Online ads from other web addresses are a good way to promote, with a click a person is directed to the target page. A good example is Google, who wins major revenue through those sponsored links that appear in the top right of the page where search results are returned.

- *Viral marketing* – Represents the spread of marketing messages via email to existing customers or a target group which are then sent by those users to their acquaintance. The advantage is given by the speed and low cost.

Commercial email marketing, also called viral, is one of the most advanced communication technologies in business, changing the way vendors interact with customers. FTC, Federal Trade Commission, in [1], defines a commercial message type email, CEMM, as any email that has as main purpose the promotion of a product or service, including the content of web pages commercially organized. For many organizations, commercial email has become a medium of communications for sales.

The power of modern sales has become impressively based on email. Many sellers can't imagine working without this tool. A research conducted by Electronic Commerce News shows that 85% of businesses use email for their communication with existing customers, 67% test the potential customers' market, 60% use email for an internal communication and 29% of the users dispose email for online orders.

The popularity of email marketing is given by advantages such as:

- lower cost than other means of communication such as direct mail marketing;

- possibility of information's distributing to a wide audience of potential customers using a list of emails;

- dispatch time of an email is very short, order of seconds or minutes, compared to a message sent by mail, which can last even several days;

- ads can be sent to a large number of subscribers to receive communicated via email with subjects that are part of the scope of their interests;

- more than half of Internet users check their emails in a normal working day;

- a precise analysis of the impact can be followed and proved to be high when done properly;

- tracking and metrics based on responses can optimize the transmission channel of emails through a continuous process of testing, using a statistical approach of the finding results;

- email marketing is green.

Email marketing offers the possibility of intercompany communication and their clients but, there are companies that send unsolicited messages, also called spam. The priority that email marketing analyses are focused is finding the threshold beyond which existing clients treat the received advertising emails as ham, legal emails, which they wish to receive, read and respond positive in line with the target set by the companies.

The difficulty of treating emails is given by the subjectivism of the customers that receive these messages.

What is practically a spam message? To answer this question is required both definitions of this type of email.

Spam, objectively and lawfully seen, is an advertisement that was not requested and not given the consent to be received. The subjective definition, however, is more restrictive and, thus, more difficult to respect. Spam is all that recipient considers to be spam.

The consequences of sending messages that may be considered spam by customers who receive them have a negative effect on electronic business.

Legally, in Romania, if you send spam you can't be punished with anything, maybe a fine if there are people with patience, time and will for claiming an email. But, technically, things get complicated.

Yahoo Mail, Gmail, Hotmail, and other such services offer the possibility of reporting those unsolicited messages, considered spam. Thus, if more people proceed in this way, future emails received by the same type as those reported will be automatically targeted spam, but not only for those users who have complained, but for all people using that mail service.

Another consequence is the possibility to block your account by the Internet provider. The blocking may occur through "black lists" which consist of spam senders and are interpreted by anti-spam filters offered default by mail servers or installed by the owners of email addresses, blocking the access of email that are considered spam.

To avoid this situation where email marketing is considered spam, it is important to take account of some recommendations such as:

- messages are sent only to those customers who have given their express consent;
- messages must be well formulated and relevant. Each mail must be unique, as each client has preferences that may differ from the others. It is important to take into account the preferences of each client as that person to handle that email as something personal;
- the address from which emails are sent must be visible, conclusive and, preferably, always the same;
- frequency of messaging is also important. Messages sent from a distance too great when the customer has an account and when he receives the first email advertising are more likely to be considered spam. Or in other circumstances, when the client is bombarded daily with messages, is not a desired approach;
- the use of specializing firm in sending messages, a company that has made a reputation with anti-spam policies formulated and implemented well. In this way, it can be avoided the blocking of these emails to different ISPs, and the rate of messages arriving in inbox, from junk mail, is higher.

3. The quality features of an e-business

Most studies conducted on the critical factors that ensure success in e-commerce focused on limited aspects of electronic business's operations.

On the other hand, Sung, in [2], has made an identification of 16 quality characteristics by applying a Likert scale on 5 levels for evaluation. The sample consisted of large firms in Korea, Japan and USA, who were questioned about the importance of each factor on the success of electronic business. This analysis did not include, however, SMEs (Small and Medium-sized Enterprises). Feidt, in 2002, also introduced these small businesses in research, [3], small innovative companies but with a high potential for development, suggesting that e-commerce is more a strategy than a technology. Table 1 shows, comparing the 2 analysis above, the influencing factors for small and medium companies, on the one hand, and large ones, on the other.

Table 1. Comparative analysis of influence factors

Nr.	Factors for SME (Feindt)	Factors for large companies (Sung)
1	Competitive profile	E-commerce strategy
2	Commitment	E-commerce expertise
3	Content	Variety of goods/services Plenty of information
4	Convenience	Easy to use
5	Control	Delivery of goods/services Payment process
6	Interaction	Customer services Customer privacy
7	Community	Customer relationship
8	Price sensibility	Competitive pricing
9	Brand image	No comparable factor
10	Partnership	No comparable factor
11	Process improvement	Evaluation of operations
12	Integration	Low cost operations
13	Facilitations	System stability, security and speed

Commitment. Studies such as [3] and [4] have identified a commitment to business growth and capacity of firms as a key factor for ensuring success. The growth motivation of the business can be seen by their willingness to adapt to market changes and by their strategies to constantly reinvent. They must be flexible and open to new ideas, technologies and opportunities. When necessary, businesses are prepared to seek strategic partners to provide appropriate technologies, financing, management expertise and contracts for a rapidly evolving of the business.

Content. The site serves as a central point for collecting and disseminating information related to marketing, sales and other functions of the organization, [5]. Web content must be accurate, informed, updated, understandable and related to users' needs. In addition, online sellers should provide an adequate range of products, accurate and relevant information about the products and services held to facilitate comparison of prices with other competitors by the customers.

Convenience. Convenience factor, similar to the degree of use, refer to the development of dedicated user interfaces, and also designs and models presented on the site, [5]. E-commerce sites can implement navigation tools, such as a menu, site map, search engines, audio, video, animation and games, tools for comparing products and services to improve communication. The problem with this is finding that balance between attractive design and fast loading time because "the system response time is inversely proportional to user's satisfaction", [5].

Interaction. Consumers expect high quality interaction with online merchants, mostly the customer service to be prompt and responsive, a non-stop help line, mailing lists and loyalty programs, and an effective system that realizes every step of the ordering process transparent to the users. Properly carried out these activities, it is proven to be very efficient to build confidence and customer loyalty, as it is presented in [3].

Control. The control factor evaluates the systems that e-business have integrated, defined in [5], to track and respond to questions and needs of users, the feedback and products returned, the control of the ordering process and credit risks controlling. Online sellers must audit the operations regular to measure and correct any dissatisfaction of the customers.

Community. This factor emphasizes the interaction between people and online organizations that allow participants to exchange information, such as forums, blogs. In [3],

it is shown that an online active community offers great opportunities for increasing the sales and productivity.

Price sensitivity. Customers expect lower prices of products and services on the Internet. A study by BizRate, [8], showed that the competitive environment forced electronic businesses to provide a reduction in transport cost, with cases in which it really is free.

Brand image. On the Internet, businesses are only one click away and customers are defined only within the language used in sites and geographical regions in which orders can be shipped concerning the economical view. Defining a brand name is essential to differentiate from other competitors in the industry. Lee and Kozar, in [7], have shown that users are willing to pay more for a company with prestige, reputation and a positive image of the brand.

Partnership. For small businesses, partnership is extremely important because of lower resources which force to work mostly closely with customers, suppliers and business partners for brand awareness, improve marketing, launching customized products. It is better to join in business community, because the position of state may attract new businesses and potential partners, being highlighted in the work [6].

Process improvement. Business process uses information technology to automate the organizational processes. This principle is easily applied to large companies, but in SMEs, financial constraints block the process.

Integration. Integrated electronic data exchange, EDI, is a system of services of quality, performance and productivity to support a large volume of business, [3]. Most small businesses don't have the capacity for full integration only if there is support from other partners in technologically advanced.

Following the studies made by Tan, [10], a correlation matrix is resulted, table 2, having 11 influencing factors previously defined. The main diagonal is equal cu 1, state that explains that any variable is all linked, 100% correlated with itself.

We define " \cap " as the correlation between two variables as:

$$x_i \cap x_i = 1$$

$$x_i \cap x_j \in [0;1]$$

For example, the strongest correlation is between Interaction and Content, with value 0.628, which means that the interaction is explained with a rate of 62.8% on web page content.

Table 2. Pearson correlation matrix

	CM	CT	CV	IR	CL	CY	PS	BI	PA	PI	SI
CM	1.000										
CT	0.540	1.000									
CV	0.191	0.494	1.000								
IR	0.422	0.628	0.299	1.000							
CL	0.137	0.339	0.478	0.454	1.000						
CY	0.494	0.350	0.464	0.359	0.361	1.000					
PS	0.192	0.182	0.175	0.219	0.087	0.310	1.000				
BI	0.286	0.251	0.261	0.254	0.187	0.336	0.233	1.000			
PA	0.207	0.269	0.471	0.159	0.470	0.510	0.236	0.179	1.000		
PI	0.220	0.420	0.549	0.334	0.388	0.462	0.243	0.319	0.423	1.000	
SI	0.227	0.356	0.554	0.394	0.537	0.432	0.167	0.273	0.435	0.378	1.000

CM: Commitment; CT: Content; CV: Convenience; IR: Interaction; CL: Control;
CY: Community; PS: Price Sensitivity; BI: Brand Image; PA: Partnership;
PI: Process Improvement; SI: Integration.

The results indicated by this analysis reveals that the 11 factors listed have a correlation between them, some more than others, correlations of which it must be taken into account in a proper marketing analysis. All parts are interconnected, not only improving a process to ensure success, it is important to find that way of comparative and correlate analysis between the main quality factors of an electronic business.

4. Metrics for analyzing email marketing's impact on customers

A good method of analyzing the email marketing's impact on customers is given by the interpretation of metrics showing in detail the main characteristics of electronic business and the level at which they stand.

Thus, research on email marketing and spam are going to 2 lines. The first is mainly focused on reducing spam from a perspective as broad as possible, and the second includes marketing studies to determine the response rate within a promotional campaign through email. For this analysis, the definitions of metrics follow, metrics that are calculated based on the results of a campaign through email but, prior to this campaign, for a further optimization at all stages of implementation.

Sent rate (*SR*) is the percentage of emails that were successfully sent in the total number of emails sent.

$$SR = \frac{\text{Number of emails successfully sent}}{\text{Number of total emails sent}}$$

, with $SR \in [0; 1]$.

If the emails are not sent for the first time to a list of addresses, this indicator should be included in the interval $[0.9; 1]$. If this indicator has values less than 0.9, this should be further examined to find the causes of this problem because such messages sent do not reach the recipients.

Opening rate (*OPR*) is the percentage of emails read by the customers of all emails sent successfully. This analysis can be performed on the total number of single openings and also on the multiple ones.

$$OPR = \frac{\text{Number of opens}}{\text{Total number of emails successfully sent}}$$

, with $OPR \in [0; 1]$.

Access rate (*AR*) is a very important metric, expressing how interested the customers were of those messages sent. It is calculated as the percentage of emails accessed, representing the following of the links in the message, to the total number of emails sent successfully.

$$AR = \frac{\text{Number of accessed email}}{\text{Total number of emails successfully sent}}$$

, with $AR \in [0; 1]$.

Unsubscribe rate (*UR*) is the percentage formed by the number of client that requested to unsubscribe from the electronic services provided through email to the total number of emails sent successfully.

$$UR = \frac{\text{Number of unsubscribes}}{\text{Total number of emails successfully sent}}$$

, with $UR \in [0; 1]$.

Forwarding rate (*FR*) is calculated if emails sent have a section that allows forwarding that message to another destination. It is calculated by the number of re-sends divided by the total number of emails successfully sent.

$$FR = \frac{\text{Number of forwarding}}{\text{Total number of emails successfully sent}}$$

, with $FR \in [0; 1]$.

Spam rate (*SPR*) is defined by the percentage of spam complains to the total number of email successfully sent.

$$SPR = \frac{\text{Number of spam complains}}{\text{Total number of emails successfully sent}}$$

, with $SPR \in [0; 1]$.

Response rate (RPR) is a metric that analyzes the last step, practical purpose of email marketing, namely the interaction between customers and e-business that is completed with the purchase of products or services.

$$RPR = \frac{\text{Number of completed orders}}{\text{Total number of emails successfully sent}}$$

, with $RPR \in [0; 1]$.

Spam probability (PSPAM) represents the weight given to the entire campaign sent emails to be classified as spam.

The main cause that occurs when an email sent is treated as spam by the recipient is explained by spam filters. In addition to the concepts defined in the preceding chapters of black lists and white lists, spam filters also have other technologies for classifying the received messages. Bayesian filtering contains an additional module which focuses on the principle of bayesian classification. This technique gives two probabilities for each word that appears in the email using the formula for conditional probabilities defined by Bayes, namely:

$$P(H/D) = \frac{P(D/H) \cdot P(H)}{P(D)}$$

, where:

- H is the hypothesis and D data;
- $P(H)$ is the prior probability of H, the probability that H is true before the data D to be seen;
- $P(D/H)$ means the conditional probability assigned to the data D and the hypothesis H being satisfied;
- $P(D)$ represents the probability of data D to be realized;
- $P(H/D)$ is the probability that hypothesis H is satisfied, data D being given in reference to the earlier assumption.

From this formula, it was calculated the Bayesian formula of probabilities that a word must belong to a spam mail as:

$$P(S/W) = \frac{P(W/S)}{P(W/M)}$$

, where:

- $P(S/W)$ is the probability that the email containing the word W to be spam;
- $P(W/S)$ represents the probability of occurrence of the word W in spam emails;
- $P(W/M)$ is the probability of occurrence of the word W in all emails held, spam and ham.

Complementary probability above formulated to total $P(T) = 1$ is described by the formula:

$$P(H/W) = \frac{P(W/H)}{P(W/M)}$$

, where:

- $P(H/W)$ is the probability that the email containing the word W is ham;
- $P(W/H)$ stands for the probability of occurrence of the word W in ham emails.

As $H + S = M$ implies that $P\left(\frac{(H + S)}{M}\right) = P\left(\frac{M}{M}\right) = P(1) = 1$.

But, $P\left(\frac{(H + S)}{M}\right) = 1 = P\left(\frac{H}{M} + \frac{S}{M}\right) = P\left(\frac{H}{M}\right) + P\left(\frac{S}{M}\right)$. It is

enough to calculate the probability of an email containing a specific word W to be spam, $P\left(\frac{S}{W}\right)$ as the complementary probability, $P\left(\frac{H}{W}\right) = 1 - P\left(\frac{S}{W}\right)$.

Having defined the formula for a word appeared in an email, to extend to the entire set of words belonging to an email, the composition al probabilities was applied, due to the independent events:

$$P(SPAM/MAIL) = P(S/W_1) * P(S/W_2) * ... * P(S/W_n)$$

, where W_i , with $i \in \{1, 2, \dots, n\}$, represents the n words belonging to the tested email. The value of n can be fixed or variable, depending on the desired implementation and optimization of the final test results.

For $PSPAM$ of all promotional campaigns, we aggregate the individual probabilities of each mail sent as:

$$PSPAM = P(SPAM/MAIL_1) * P(SPAM/MAIL_2) * ... * P(SPAM/MAIL_m)$$

, where:

- $MAIL_i$, with $i \in \{1, 2, \dots, m\}$, represents the promotional emails and m the number of emails sent.

5. Defining an optimization model in an email campaign

To implement the principles presented in the previous chapters and of the analysis of the indicators defined, a progressive model is realized, consisted of steps to follow to improve quality results in an ad campaign run by email communication channel.

For this, there is need of a previous collection of data with reference to existing customers database containing information on name, email, address, age, gender, occupation and categories of interest of those customers as being buyers of products and services owned and marketed by the e-business. This information will be collected from the customers when they create accounts through the web site and integrated into a well structured database.

The following is a breakdown of each step containing figure 1, along with summarizing the statistical principles for a better understanding of the concept of statistical survey.

Step1. Verification of customers' database. As stated, the database must be constantly updated, for existing no differences between this information and the reality..

Step2. Determination of survey's dimension. To determine the sample size is required to conduct a statistical analysis of the customers. Next, we consider the statistical population as the entire database of the customers.

Suppose we want to find the dispersion of the population when it is divided by gender. For a simple random selection of non-renewable, and a limit error chosen with a probability guarantee of results, sample size, n , is directly proportional to the dispersion of population and inversely proportional to the size of the limit error. Proposed sample size will be equal to:

$$n = \frac{z^2 w(1-w)}{\Delta_w^2}$$

, where:

- z^2 is the probability indicator that is chosen from statistical tables for a 1- α confidence level chosen;
- $w(1-w)$ is the dispersion of the population with the gender classification, variable with two states, 1 or 0, w means the weight of the population that is 0, female;
- Δ_w refers to the limit error chosen, expressed as a percentage.

Knowing the sample size, it remains to choose those n customers from the total population to join the next survey. Choice may be left to the program, choosing randomly, or can be generated in a different way of sorting. To simplify the process, will choose the random method.

Step3. Training the spam filter. Bayesian spam filter is a type of artificial intelligence method, one of the initiators in this field being Paul Graham in [12], [13] and [14], a method involving previously training. Training is conducted on the basis of emails collected from various sources in the company, to be able to simulate as conclusive as possible a spam filter that a customer might use it. After all emails are loaded, the filter will calculate a spam weight for each word encountered in the emails. This weight will be used in the step where PSPAM probability will be calculated.

Step4. Formulation of advertising emails. This phase includes the formulation, analysis in terms of social, psychological, economic and visual future advertising emails from the promotional campaign. It will be taken in account of previous results made, for a better understanding of the customers, their way of thinking and decision making.

Step5. PSPAM calculation. Before sending the emails, PSPAM indicator is calculated as defined in the previous chapter in order to make an earlier analysis of the sent messages. The purpose is to allow internal simulation of the messages' interpretations by the customers.

Step6. Probability's analysis. The metric PSPAM is compared with a threshold given, suppose the value of 0.1. If the metric's value is below this threshold means that the emails have the probability to be considered spam less than 1%, and thus, step 7 will be preceded. Otherwise, the messages do not pass the first stage of analysis and will be redrafted and returned to step 4.

Step7. Emails' sending to the sample set. The emails are sent to the customers selected in step 2.

Step8. Analysis of the sample's results. Based on the results received and the feedback provided by sending the n emails, the metrics defined in chapter 4 are calculated.

Step9. Decision the extend the sample. The metrics' results are economical analyzed and a decision is being taken whether it can continue to expand the sample throughout the population. If the results are favorable, step 10 is preceded. In the other case, the process returns to step 4, for an reassessment of the existing situation.

Step10. Sending the remaining emails. At this stage, the advertising emails are considered to be as close to that level desired by the company, having a high probability of success for the entire set of customers. The remaining emails are sent to the rest of the customers.

Step11. Final analysis. The process being finished, the last conclusions are drawn reanalyzing the defined metrics, optimizing the conclusions for a future campaign correction.

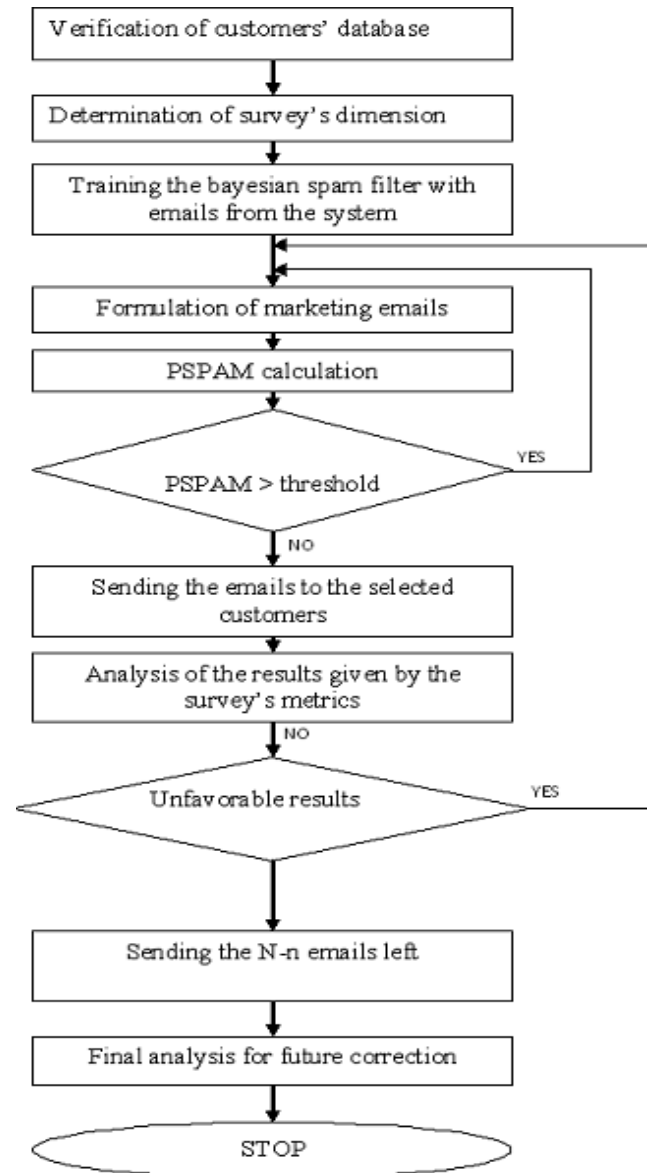


Fig 1. Steps of email promoting model's implementation

The principle of survey was chosen because it is more convenient for an electronic business to test an ad campaign on a more limited range of customers than to send a message that will be interpreted negatively to a much larger population. This theory is based on the trust that given by the customers to online businesses and the level of acceptance of entering into their privacy.

6. Conclusions

The knowledge society, along with the Internet communication channel, particularly the email, has made the promoting ways of businesses to get a different structure. Electronic businesses are seen by grown demanding of the customers who expect accurate information, concrete, rapid information to be made in a honest, convenient, interactive, to meet all the needs that a customer has.

Marketing's new paradigm, described in [15], paradigm which is introduced by the web, makes a change from product to customer, this including a new personalized micro-level and a management of customer relationship. The problems that arise in such a

framework are given for how such information should be submitted, to comply with customer's privacy, without diminishing the intensity of products and online services' promoting

The model presented in chapter 5 optimizes the flow of information and focuses the tests made for market evaluation on a limited number of customers. The major benefit is reducing the negative effects of an advertising campaign that hasn't the correct target group chosen, doesn't transmit the desired and expected information to the customers and doesn't consider the preference generally provided by the users when registering in the internal e-business system.

The best way to see the difference between email marketing and spam is in terms of the customers, those who receive these messages, because, as it was presented, spam is something subjective, in addition to legal and objective vision given by various laws, research and interpretations.

Spam is all that the recipient considers to be spam, so all we have to do left is to model the email sent to customer requirements, to personalize and optimize, to transform it from an advertising email to a personal one, applying sociological, mental, statistical, artificial intelligence and marketing theories.

References

1. **Federal Trade Commission** <http://www.ftc.gov/>
2. Sung, T. K. **E-commerce critical success factors: East vs. West**, Technological Forecasting and Social Change, Vol. 73, ISSN 0040-1625, 2006, pp. 1161-1177
3. Feindt, S., Jeffcoate, J. and Chappell, C. **Identifying success factors for rapid growth in SME e-commerce**, Small Business Economics Review, Vol. 19, ISSN 0921-898, 2002, pp. 51-62
4. Taylor, M. and Murphy, A. **SMEs and e-business**, Journal of Small Business and Enterprise Development, Vol. 11, No. 3, ISSN 1462-6004, 2004, pp. 280-289
5. Rao, S. S., Metts, G. and Monge, C. A. M. **Electronic commerce development in small and medium enterprises: a stage model and its implications**, Business Process Management Journal, Vol. 9, ISSN 1463-7154, 2003, pp. 11-32
6. Jeffcoate, J., Chappell, C. and Feindt, S. **Best practice in SME adoption of e-commerce**, Benchmarking: An International Journal, Vol. 9, No. 2, ISSN 1710-2251, 2002, pp. 122-132
7. Lee, Y. and Kozar, K. A. **Investigating the effect of website quality on e-business success: an Analytic Hierarchical Approach (AHP) approach**, Decision Support Systems, Vol. 42, ISSN 0167-9236, 2006, pp. 1383-1401
8. BizRate **Best practices and tips for a successful holiday in 2006, e Holiday Mood Study**, BizRate Research, <http://www.shop.org/eholiday06/download/eholidayreport.pdf>
9. Poon, S. **Business environment and internet commerce benefit: a small business perspective**, European Journal of Information System, Vol. 9, ISSN 0960-085X, 2000, pp. 72-81
10. Tan, P. J. S., Sharma, R. S. and Theng, Y. L. **Effective e-commerce strategies for small online retailers**, International Journal of Electronic Business, Vol. 7, No. 5, ISSN 1470-6067, 2009, pp. 445-472
11. Poon, S. and Swatman, P.M.C. **An exploratory study of small business internet commerce issues**, Information and Management, Vol. 35, ISSN 0378-7206, 1999, pp. 9-18,
12. Worthz, J., Graham, N. and Sapte, D.W. **New rules for electronic marketing – An obstacle to m-commerce?**, Computer Law and Security Report, Vol. 18, No. 2, ISSN 0267-3649, 2002

13. Graham, P. **Better Bayesian filtering**, January 2003, <http://paulgraham.com/better.html>
14. Graham, P. **A plan for Spam**, August 2002, <http://paulgraham.com/spam.html>
15. Rust, R. T. and Espinoza, F. **How technology advances influence business research and marketing strategy**, Journal Business Resource, No. 59, ISSN 1537-8020, 2006, pp. 1072-1078

¹ **Ion SMEUREANU** has graduated the Faculty of Planning and Economic Cybernetics in 1980, as promotion leader. He holds a PhD diploma in "Economic Cybernetics" from 1992 and has a remarkable didactic activity since 1984, when he joined the staff of Bucharest Academy of Economic Studies. Currently, he is a full Professor of Economic Informatics within the Department of Economic Informatics and the dean of the Faculty of Cybernetics, Statistics and Economic Informatics from the Bucharest Academy of Economic Studies. He is the author of more than 16 books and an impressive number of articles. He was also project coordinator or member in many important research projects. He was awarded the Nicolae Georgescu-Roegen diploma, the award for the entire research activity offered by the Romanian Statistics Society in 2007 and many others.

² **Madalina ZURINI** has graduated the Faculty of Cybernetics, Statistics and Economic Informatics in 2008 and she is currently conducting a master in computer science researching the *implications of Bayesian classifications for optimizing spam filters*. She is also engaged in a Pedagogical Program as part of the Department of Pedagogical Studies. Her fields of interest are data classification, artificial intelligence, algorithm optimisations, mathematical and econometrical analysis of factors' influences. She wants to pursue a pedagogical career and also to give her PhD in the same area of research