A Posterior Distribution for Anti-spam Bayesian Statistical Model

Youcef Begriche,
IEEE member
Paris,France
Email: Youcef.Begriche@ieee.org

Houda Labiod,
Institut Telecom, TelecomParisTech, LTCI-UMR 5141 CNRS
46 rue Barrault 75013 Paris France
Email: Houda.Labiod@telecom-paristech.fr

***Abstract*—This paper deals with Bayesian models applied to anti-spam. In our previous work involved in mathematically modeling the behavior of spam, we worked on defining the priors distributions that summarize the behavior of spam. In this paper, taking into account certain priors distributions and we define the posterior distributions that summarize the probability that a received message is spam. These laws a priori and a posteriori certainly help improve the spam detection and increase the reliability decision.**

Keywords - Beta distribution, Normal distribution, Conditional density, prior law, posterior law, Spam, Ham, Bayesian statistical model, transfer theorem, Classification.

I. INTRODUCTION

This article aims to contribute to the study of anti-spam methods by providing a statistical Bayesian model built on two laws: prior and posterior. The negative effects of spam on information systems were widely described in [1]. Several research works and efforts are being made to reduce their scale and impact. Though not satisfied with this phenomenon, the community has nevertheless been taking into account its existence for designing network architectures and security implementation. Solutions have been developed but are still not sufficient for a total eradication. These solutions are based on several methods of treatment, and can be found in [2]. Many of these methods require a specific network infrastructure or cooperation between network elements (e.g. mail server directories) based on particular organizations that require specific costs. Anti-spam Bayesian methods operate as "standalone" and are based on statistical and probabilistic mathematical models. These methods benefit from an important research in the field of language processing. Indeed, fundamental results in this field have been obtained [3] and make it possible to apply it to anti-spam field. Paul Graham [4] was probably a pioneer in suggesting a complete solution to filter messages based on Bayesian statistical methods [5,6]. Several classification-based solutions also exist for which a summary can be found in [2]. The author proposes three main classes :

- Filtering the content amounts to distinguish the spam from the non spam (called ham). As mentioned in [2], this type of process is prone to errors, and subject to false negatives [1] and false positives [2]. The goal is to reduce the number of false positives. For this purpose there are several filtering methods: heuristic, adaptive or Bayesian filtering, collaborative and honey pot.
- Identifying one or more fields of the message. Classical cryptographic functions may be used (digital signature and authentication). This method is effective against phishing [3] or any approach of spoofing.
- Establishing a cost for message transmission. This method reduces a significant amount of resources at spammer side. Several approaches based on costs exist.

Because of their limitations when they are used separately, administrators and users sometimes combine several of these methods.

In [7], since $P(C = spam)$ and $P(C = ham)$ are estimated by dividing the number of training messages of category (spam or ham) by the total number of training messages, authors consider the probability having a spam $P(C = spam)$ as a random variable. Based on Jeffrey's law and Fisher's information they determine its law which is a $Beta(0.5, 0.5)$. Our work in this paper extends the result obtained in [4] focusing on the posterior law.

In this paper, we consider the probability $p(\overrightarrow{x}/s)$ as a random variable. Using the transfer theorem, we give its law from that of p, depending on the parameter c. Graham [4] and Robinson [8] worked on messages content by assuming the probability of a message being a spam to be equal to that of a ham ($P(spam) = P(ham) = 0.5$). By considering a prior law on the probability of a message being a spam, we obtained the corresponding law after observing the parameter c in a message; that is the posterior law.

In [10], some Naive Bayes (NB) classifiers are proposed : multi-variate Bernoulli NB, multinomial NB, Boolean attributes, multi-variate Gauss NB.

The classifiers are built on Bayes' theorem: if a message is represented by a vector $\overrightarrow{x} = (x_1, \ldots, x_n)$, then the probability that the message belongs to the class c (spam or ham) is given by the formula

$$P(c/\overrightarrow{x}) = \frac{P(c).P(\overrightarrow{x}/c)}{P(\overrightarrow{x})} \tag{1}$$

As there are two classes (spam and ham) the denominator is written

$$P(\overrightarrow{x}) = P(\overrightarrow{x}/s).P(s) + P(\overrightarrow{x}/h).P(h) \tag{2}$$

978-1-4577-0737-7/11/$26.00 ©2011 IEEE

In this paper, we work on Formulas (1) and (2) which contain 4 parameters : $P(s)$, $P(h) = 1 - P(s)$, $P(\overrightarrow{x}/s)$ and $P(\overrightarrow{x}/h)$. Given the relationship between them we reduce that number to 2. We also take into account the results obtained in [4], *i.e.* $P(s)$ is a random variable
The criterion for classifying a message as spam is given by

$$\frac{P(s).P(\overrightarrow{x}/s)}{P(\overrightarrow{x}/s).P(s) + P(\overrightarrow{x}/h).P(h)} > T \qquad (3)$$

where T is a fixed threshold.
The criterion for classifying is adapted for each classifier:

1) For the multivariate Bernoulli NB distribution, the criterion becomes

$$\frac{P(s).\prod_{i=1}^{m} P(t_i/s)^{x_i}.(1-P(t_i/s))^{(1-x_i)}}{\sum_{c\in\{s,h\}} P(c).\prod_{i=1}^{m} P(t_i/c)^{x_i}.(1-P(t_i/c))^{(1-x_i)}} > T$$

   where
   - $x_i = 1$ if the token t_i occurs in the message, otherwise $x_i = 0$.
   - $p(t/c) = \frac{1+M_{t,c}}{2+M_c}$ where $M_{t,c}$ is the number of training messages of category c that contain token t, while M_c is the total number of training messages of category c.

2) For the multinomial NB distribution, (3) becomes

$$\frac{P(s).\prod_{i=1}^{m} P(t_i/s)^{x_i}}{\sum_{c\in\{s,h\}} P(c).\prod_{i=1}^{m} P(t_i/c)^{x_i}} > T$$

3) For the multivariate Gaussian NB distribution, (3) gives

$$\frac{P(s).\prod_{i=1}^{m} g(x_i;\mu_{i,s},\sigma_{i,s})}{\sum_{c\in\{s,h\}} P(c).\prod_{i=1}^{m} g(x_i;\mu_{i,s},\sigma_{i,s})} > T$$

   where

$$g(x_i;\mu_{i,s},\sigma_{i,s}) = \frac{1}{\sigma_{i,s}\sqrt{2\Pi}} e^{-\frac{(x_i-\mu_{i,s})^2}{2\sigma_{i,s}^2}}$$

   and the mean $(\mu_{i,s})$ and standard deviation $(\sigma_{i,s})$ of each distribution are estimated from the training data.

In those classifiers, $P(s)$ and $P(h)$ are typically estimated by dividing the number of training messages of category (s or h) by the total number of training messages.
In [11], the criterion for classifying a message as a spam is

$$\frac{P(C = spam/\overrightarrow{X} = \overrightarrow{x})}{P(C = ham/\overrightarrow{X} = \overrightarrow{x})} > \lambda \qquad (4)$$

where

$$P(C = c/\overrightarrow{X} = \overrightarrow{x}) = \frac{P(C = c)\prod_{i=1}^{n} P(X_i = x_i/C = c)}{\sum_{k\in\{spam,ham\}} P(C = k)\prod_{i=1}^{n} P(X_i = x_i/C = k)}$$

where n is the size of the vector $\overrightarrow{x}$.
For this classifier, $P(C = spam)$ and $P(C = ham)$ are also estimated by dividing the number of training messages of category (spam or ham) by the total number of training messages.
In [12], authors use a similar formulae to (1) and (2) as follows

$$\begin{aligned} p(s/\overrightarrow{x}) &= \frac{P(s).p(\overrightarrow{x}/s)}{P(\overrightarrow{x}/s).P(s) + P(\overrightarrow{x}/h).P(h)} \\ &= \frac{\frac{P(\overrightarrow{x}/s)}{P(\overrightarrow{x}/h)} P(s)}{\frac{P(\overrightarrow{x}/s)}{P(\overrightarrow{x}/h)} P(s) + P(h)}. \end{aligned} \qquad (5)$$

In this paper, we consider only the ratio $\frac{P(\overrightarrow{x}/s)}{P(\overrightarrow{x}/h)}$. $P(C = spam)$ and $P(C = ham)$ are also estimated by dividing the number of training messages of category (spam or ham) by the total number of training messages.
In [13], the following criterion is used :
Given an input document d, its target class (spam, ham) can be found by choosing the one with the high posterior probability.

$$H(d) = \arg\max_{c_j\in C} \sum_{w\in F} \frac{P(w/c_j)P(c_j)}{\sum_{c'\in C} P(w/c')P(c')} P(w/d')CHI(w,c_j)$$

where F is a text feature vector; $P(w/c_j)$ is the ratio of frequency of word w in d among the total number of words, $CHI(w,c_j)$ is a χ^2 statistic of word w and class c_j which measures the dependence between word w and c_j.
$P(c_j)$ is the number of documents with class label c_j divided by the total number of documents.
In [14], a simple linear regression model is considered. This model $Y_M = a + bR_M + \varepsilon_M$ relates the class label to the rank of the message, where Y_M is a dichotomous outcome of the M^{th} message (spam or ham), R_M is the rank of the M^{th} message and ε_M is a random error. Assuming that $E(\varepsilon_M) = 0$, *i.e.* no random error on the average, then $E(Y_M) = a + bR_M$. Let $p(Y_M = spam) = \Pi_M$ be the probability that the M^{th} message is a spam. They took

$$p(Y_M = spam) = \frac{e^{(a+bR_M)}}{1 + e^{(a+bR_M)}}.$$

In [15], the maximum a posterior (MAP) class is obtained by calculating

$$e_{MAP} = \arg\max_{e_i\in E} P(e_i)\prod p(k_j/e_i)$$

where k_j is the word found in the j^{th} position in the unseen query. e_i is the tutorial example/page.

$$P(e_i) = \frac{Number\ of\ queries\ in\ Q\ with\ e_i}{Number\ of\ queries\ in\ Q}$$

where Q is the queries set.

In [16,17,18], the degree of confidence of $\overrightarrow{x}$ being a spam is introduced:

$$\begin{aligned} W_S^{NB}(\overrightarrow{x}) &:= P(spam/\overrightarrow{x}) \\ &= \frac{p(spam)\cdot\prod_{i=1}^{m}P(x_i/spam)}{\sum_{k\in\{spam,ham\}}P(k)\cdot\prod_{i=1}^{m}P(x_i/k)} \end{aligned}$$

and $P(C = spam)$ and $P(C = ham)$ are also estimated by dividing the number of training messages of category (spam or legitime) by the total number of training messages.

In [19], authors deliver a discussion about the implementation of Binomial and Poisson distribution in Bayesian spam filter, to calculate the probability of a mail being a spam, containing words that are not already stored in a database (*i.e.*, encountered by the filter for the first time). They use :

- A Binomial random variable R with parameters n and p, then the probability function of R is given as follows :

  $$f(r) = P(R = r) = \binom{n}{k}p^r q^{n-r}, \quad r = 0, 1, 2, 3, \ldots, n$$

  where $p + q = 1$.
- A Poisson random variable R, with probability function

  $$f(r) = P(R = r) = e^{-m}m^r/r!, \quad r = 0, 1, 2, 3, \ldots, n$$

  where $m = np$.

p is the probability to have a spam word, which is assumed to be constant from trial to trial (0.4 for the Binomial distribution and 0.04 for the Poisson distribution).

In [20], author defines a local probability. For example:

$$P_{local-spam} = 0.5 + \frac{(N_{spam} - N_{nonspam})}{C_1 \times (N_{spam} + N_{nonspam} + C_2)}.$$

Each word feature generates one such local probability. These local probabilities are then used in a Bayesian chain rule to calculate an overall probability that an incoming text is a spam.

The Bayesian chain rule is
$P(in\ class/feature) =$

$$\frac{P(feat/inclass)\times P(in\ class)}{P(feat/in\ class)\times P(in\ class)+P(feat/not\ in\ class)\times P(not\ in\ class)}$$

which is applied iteratively to each local chain probability feature into an overall probability for the text. Here $p(in\ class)$ is also estimated by dividing the number of training messages of class (spam or ham) by the total number of training messages.

In [21], they use the formula

$$\begin{aligned} y &= \log\left[\frac{P(c_0/\overrightarrow{x})}{P(c_1/\overrightarrow{x})}\right] \\ &= \log[P(c_0)] - \log[P(c_1)] + \sum_{i=1}^{n} w_i \end{aligned}$$

where $w_i = \log\left[\frac{P(x_i/c_0)}{P(x_i/c_1)}\right]$
and calculate

$$P(c_0/\overrightarrow{x}) = \frac{1}{1 + 2^{-y}}.$$

$P(C = spam)$ and $P(C = ham)$ are also estimated by dividing the number of training messages of category (spam or ham) by the total number of training messages.

In [22], they introduce a score notion as follows. If an e-mail is represented by $\overrightarrow{x} = (x_1, \ldots, x_n)$ then

$$\begin{aligned} score(\overrightarrow{x}) &= \log[P(spam)] + \sum_k \log[P(x_k/spam)] \\ &- \log[P(leg)] - \sum_k \log[P(x_k/leg)]. \end{aligned}$$

Therefore, if $score(\overrightarrow{x}) > 0$, the e-mail will be a spam, and legitime otherwise.

In this work, we complete the work done in [4] by proposing a posterior law for anti-spam using Formula (1). So, we give the posterior law which completes the statistical model started in [21] modeling the evolution of spam.

This work differs from the previous works listed above; it does not assume that the probability of having a spam is equal to 0.5. It's the following of work done in [23] where they consider that spam law varies following beta law, in the light of this consideration it gives the nature of the posterior law.

We introduce a parameter c which is the ratio $\frac{P(\overrightarrow{x}/s)}{P(\overrightarrow{x}/h)}$, and we control the probability $P(c/\overrightarrow{x})$ depending on the parameter c.

To ensure the accuracy of this work, we repeat the same work with the normal law as *a priori*, find the posterior law and at the end, we compare the results between them and we give some conclusions.

II. TRANSFER THEOREM[24]

In this section we present the transfer theorem which will be used in our method.

THEOREM:

Let X be a random variable taking its values in R^d, f its probability density, and let φ a Borelian function from R^d to R (*i.e.* measurable function).

Then if one of the terms of the equality $E[\varphi(X)] = \int_{R^d}\varphi(u)f(u)du$ has a meaning, then the other also does, and the equality occurs.

Conversely, if the above equality takes place for any bounded Borelian function φ then f is a probability density of X.

EXAMPLE:

If X is a random variable with standard normal distribution, *i.e.* its density is $f_X = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ for all real x.

Using the transfer theorem we can find the law of the random variable X^2.

Let φ a bounded and continuous function. Then, by a change of variable

$$\begin{aligned} E[\varphi(X^2)] &= \int_{-\infty}^{+\infty} \varphi(x^2)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \\ &= \int_{0}^{+\infty} \varphi(u)[\frac{1}{\sqrt{2\pi}}u^{-\frac{1}{2}}e^{-\frac{u}{2}}]du. \end{aligned}$$

Thus, the density of X^2 is :

$$\begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{\sqrt{2\Pi}}u^{-\frac{1}{2}}e^{-\frac{u}{2}} & \text{if } x \geq 0 \end{cases}$$

A. *Calculation of the posterior law for beta law*

Using Formulae (1) (with $c = s$) and (2) we get

$$\begin{aligned} P(s/\overrightarrow{x}) &= \frac{1}{1+\frac{P(\overrightarrow{x}/h)}{P(\overrightarrow{x}/s)}\times\frac{P(h)}{P(s)}} \\ &= \frac{1}{1+\frac{P(\overrightarrow{x}/h)}{P(\overrightarrow{x}/s)}\times(\frac{1}{P(s)}-1)} \end{aligned}$$

by $P(h) = 1 - P(s)$.
From [4], $P(s)$ is a random variable with beta(0.5;0.5) law and noted X.

The ratio $\frac{P(\overrightarrow{x}/s)}{P(\overrightarrow{x}/h)}$ is noted c.

Thus, the aim is to seek the law of probability of the random variable

$$Y = \frac{1}{1+\frac{1}{c}\times(\frac{1}{X}-1)} = \frac{X}{(1-\frac{1}{c})X+\frac{1}{c}} \qquad (6)$$

where X is a beta-distributed random variable with density f_X and c is a real parameter.
In the following, we use the transfer theorem. Let φ a measurable function, then

$$\begin{aligned} E[\varphi(Y)] &= \int_{-\infty}^{+\infty}\varphi\left[\frac{x}{(1-\frac{1}{c})x+\frac{1}{c}}\right]f_X(x)dx \qquad (7) \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1\varphi\left[\frac{x}{(1-\frac{1}{c})x+\frac{1}{c}}\right]x^{\alpha-1}(1-x)^{\beta-1}dx \end{aligned}$$

Using the change of variable

$$y = \frac{x}{x(1-\frac{1}{c})+\frac{1}{c}}, \qquad (8)$$

we get

$$E[\varphi(Y)] = \int_{-\infty}^{+\infty}\varphi(y)f_Y(y)dy$$

where

$$f_Y(y) = \frac{1}{c}\times\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{(\frac{1}{c}y)^{\alpha-1}(1-y)^{\beta-1}}{[1-y(1-\frac{1}{c})]^{\alpha+\beta}}\cdot 1_{[0,1]}(y)$$

is the probability density function of Y, according to the transfer theorem.
For $\alpha = \beta = 0.5$, we have $\Gamma(\alpha) = \Gamma(\beta) = \sqrt{\pi}$ and $\Gamma(\alpha+\beta) = \Gamma(1) = 1$.
The posterior density of Y is then

$$f_Y(y) = -\frac{\sqrt{c}}{\pi}\cdot\frac{1}{\sqrt{y(1-y)}\cdot[(c-1)y-c]}\cdot 1_{[0,1]}(y)$$

and its distribution function is

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} -\frac{\sqrt{c}}{\pi}\cdot\frac{1}{\sqrt{t(1-t)}\cdot[(c-1)t-c]}\cdot 1_{[0,1]}(t)dt \\ &= \begin{cases} 0 & \text{if } y \leq 0, \\ \int_0^y -\frac{\sqrt{c}}{\pi}\cdot\frac{1}{\sqrt{t(1-t)}\cdot[(c-1)t-c]}dt & \text{if } 0 < y < 1, \\ 1 & \text{if } y \geq 1. \end{cases} \end{aligned}$$

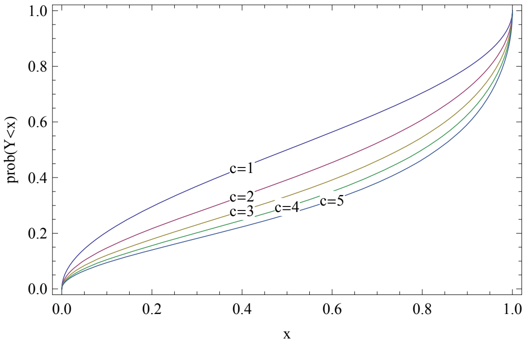

Fig. 1. Distribution functions for posterior law

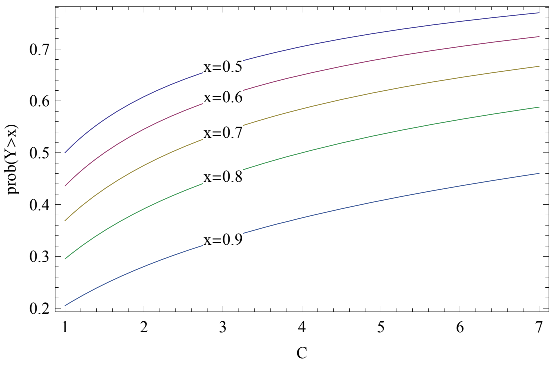

Fig. 2. Probability that Y is greater or equal than y as a function of c, for Beta distribution

B. *Calculation of the posterior law for uniforme law*

We consider the case where the random variable X follows uniforme on [0;1] distribution. The density function is then:

$$f_X(x) = 1_{[0,1]}(x)$$

In order to get the law of the random variable Y defined by (6) we use the same steps as in (7) and (8) and we get

$$E[\varphi(Y)] = \int_{-\infty}^{+\infty}\varphi(y)\cdot f_Y(y)dy$$

where

$$f_Y(y) = \frac{1}{c\left[y(1-\frac{1}{c})-1\right]^2}\cdot 1_{[0,1]}(y)$$

is the density of Y, by the transfer theorem.
and its distribution function is

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{y} \frac{1}{c\left[t(1-\frac{1}{c})-1\right]^2}\cdot 1_{[0,1]}(t)dt \\ &= \begin{cases} 0 & \text{if } y \leq 0, \\ \frac{-y}{c((c-1)y-c)} & \text{if } 0 < y < 1, \\ 1 & \text{if } y \geq 1. \end{cases} \end{aligned}$$

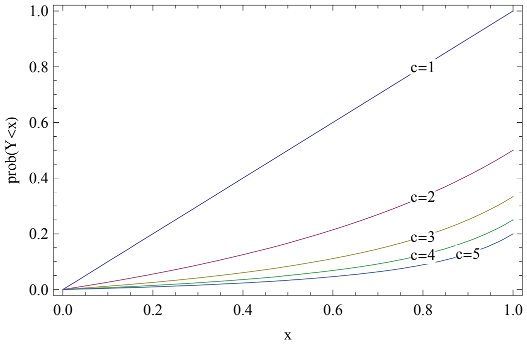

Fig. 3. Distribution functions for posterior law

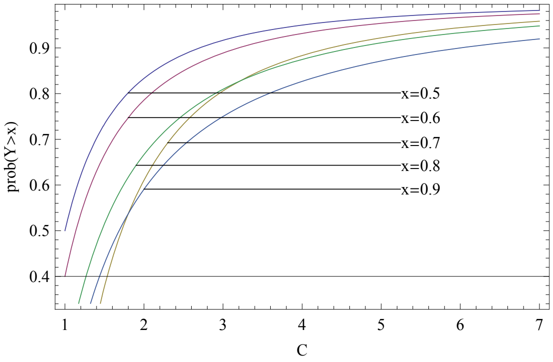

Fig. 4. Probability that Y is greater or equal than x as a function of c, for uniforme distribution

III. Results for the posterior law of beta law

In this section, we provide the values of the probability $P(Y \geq y)$ for some values of c.

Figure 2 shows the curves of $P(Y \geq y)$, which is the probability of the spam's acceptance area, as a function of parameter c .

In the following we try to be more specific in giving the value of the probability that Y belongs to an interval $[a, b]$, the probability $P(a \leq Y \leq y)$, for some values of c.

c	$P(Y \geq 0.5)$	$P(Y \geq 0.6)$	$P(Y \geq 0.7)$	$P(Y \geq 0.8)$
0.5	0.391827	0.333333	0.275999	0.216347
1	0.5	0.435906	0.36901	0.295167
1.5	0.564094	0.5	0.430245	0.349802
2	0.608173	0.545629	0.47549	0.391827
2.5	0.640983	0.580431	0.510978	0.425876
3	0.666667	0.608173	0.539893	0.454371

TABLE I
Probability that Y is greater than y

c	[0.5-0.6]	[0.6-0.7]	[0.7-0.8]	[0.8-0.9]
0.5	0.0625445	0.0701385	0.0836639	0.111389
1	0.0640942	0.0668957	0.0738429	0.0903345
1.5	0.0615717	0.0618285	0.0657539	0.0775833
2	0.0584932	0.0573345	0.0596519	0.0689838
2.5	0.0555534	0.0535583	0.0596519	0.0627079
3	0.0528955	0.0503843	0.0511411	0.1594

TABLE II
Probability that Y is in $[a, b]$ for Beta distribution

Table II shows the probability of the acceptance area of the spam in interval $[0.5 - 0.6]$, $[0.6 - 0.7]$, $[0.7 - 0.8]$ and $[0.8 - 0.9]$.

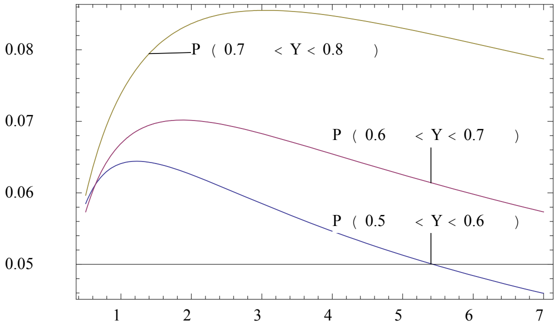

Fig. 5. Probability that Y is in $[a, b]$ as function of c, for Beta distribution

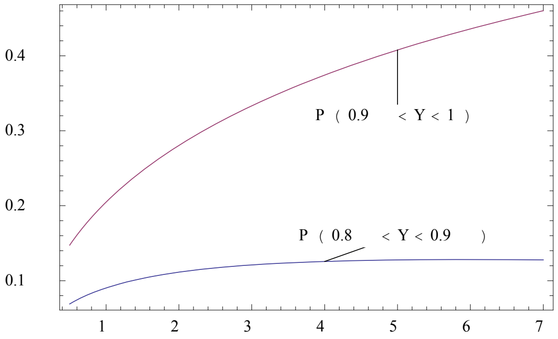

Fig. 6. Probability that Y is in [0.8,0.9] or [0.9,1] for Beta distribution

IV. Conclusion and Future Work

In this paper, we have shown using the transfer theorem that the law which summarize the initial information (beta law) built on the Fisher information is well suited to define a statistical model containing a prior law and posterior law which gives better results than the calculations based only on static case.

For future work, we intend to study a new method to detect spam considering the occurrence number of some words appearing in a message. Other prior distributions will be tested against the beta law.

References

[1] D. W. K. Khong, *An Economic Analysis of Spam Law*, Erasmus Law & Economics Review, Vol. 1, pp. 23-45, February 2004.

[2] A. Herzberg, *Controlling Spam by Secure Internet Content Selection*, Proceedings of Secure Communication Networks (SCN) 2004, LNCS vol. 3352, Ed. Springer-Verlag.

[3] M. Sahami, M. Dumais, S. Heckerman, E. Horvitz, *A Bayesian Approch to Filtering Junk E-mail*, learning for text Categorization -AAAI Workshop, pp.55-62, 1998, Madison Wisconsin.

[4] P. Graham, *A Plan for Spam*, http:// paulgraham.com/. Youcef Begriche Houda Labiod, *A Prior Distribution for Anti-spam Statistical Bayesian Model,* IFIP Network and Service Security, Paris, 2009.

[5] C. P. Robert, *Le choix Baysien. Principes et pratiques*, Ed. Springer, 2006.

[6] JJ Droesbeke, J. Fine, G. Saporta. *Mthodes Baysiennes en statistique*, Ed. Technip, 2002.

[7] Youcef Begriche Houda Labiod, *A Prior Distribution for Anti-spam Statistical Bayesian Model,* IFIP Network and Service Security, Paris, 2009.

[8] Gary Robinson, *A Statistical Approch to the Spam Problem*, Linux journal, 01-03-2003, Issue 107,2003.

[9] Y Begriche, *Modèles Bayésien Anti-Spam : principes,* CRISIS in IEEE GIIS conference, July 2007, Marrakech, Maroc.

[10] V. Metsis, I. Androutsopoulos and G. Paliouras, *Spam Filtering with Naive Bayes – Which Naive Bayes?* Proceedings of the 3rd Conference on Email and Anti-Spam (CEAS 2006), Mountain View, CA, USA, 2006.

[11] I. Androutsopoulos, G. Paliouras, V. Karkaletsis, G. Sakkis, C.D. Spyropoulos, P. Stamatopoulos , *Learning to filter Spam E-Mail: A Comparison of a Nave Bayesian and a Memory-Based Approach,* Proc. of the workshop on Machine Learning and Textual Information Access, 4th European Conference on Principles and Practice of Knowledge Discovery in Databases, 2000,France.

[12] Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim, Sung Hyon Myaeng, *Some Effective Techniques for Naive Bayes Text Classification,* IEEE Transactions on Knowledge and Data Engineering, Volume 18 Issue 11, November 2006.

[13] Yanhui Guo; Yaolong Zhang; Jianyi Liu; Cong Wang; *Research on the Comprehensive Anti-Spam Filter,* Industrial Informatics, 2006 IEEE International Conference on 16-18 Aug. 2006 Page(s):1069 - 1074

[14] Yan Zhou, Madhuri S. Mulekar, Praveen Nerellapalli: *Adaptive Spam Filtering Using Dynamic Feature Spaces,* International Journal on Artificial Intelligence Tools 16(4): 627-646 (2007)

[15] Hernes, O.; Jianna Zhang; *A tutorial search engine based on Bayesian learning,* Machine Learning and Applications, 2004. Proceedings. 2004 International Conference on 16-18 December, 2004 Page(s):418 - 422

[16] G. Sakkis, I. Androutsopoulos, G. Paliouras, V. Karkaletsis, C. D. Spyropoulos, P. Stamatopoulos *Stacking classifiers for anti-spam filtering of e-mail,* Proceedings of "Empirical Methods in Natural Language Processing" (EMNLP 2001), L. Lee and D. Harman (Eds.), pp. 44-50, Carnegie Mellon University, Pittsburgh, PA, 2001

[17] Georgios Sakkis1 , Ion Androutsopoulos2 , Georgios Paliouras1 , Vangelis Karkaletsis1 , Constantine D. Spyropoulos1 and Panagiotis Stamatopoulos2 *A Memory-Based Approach to Anti-Spam Filtering for Mailing Lists,* Information Retrieval, Volume 6, Number 1, Pages 49-73, / janvier 2003, Kluwer Academic Publishers Hingham, MA,USA,diteur Springer Netherlands

[18] I. Androutsopoulos, G. Paliouras and E. Michelakis, *Learning to Filter Unsolicited Commercial E-Mail,* Technical report 2004/2, NCSR "Demokritos", revised version (October 2004), with additional minor corrections (October 2006).

[19] Redwan Zakariah , Samina Ehsan, *Detecting junk Mails by Implementing Statistical Theory,* 20th International Conference on Advanced Information Networking and Applications (AINA 2006), 18-20 April 2006, Vienna, Austria. IEEE Computer Society 2006.

[20] Yerazunis, W.S., *"The Spam-Filtering Accuracy Plateau at 99.9% Accuracy and How to Get Past It",* MIT Spam Conference, January 2004

[21] Marsono, M.N.; El-Kharashi, M.W.; Gebali, F.; Sudhakar Ganti; *Distributed Layer-3 E-Mail Classification for Spam Control,* Electrical and Computer Engineering, 2006. CCECE '06. Canadian Conference on May 2006 Page(s):742 - 745.

[22] Zhen Yang Xiangfei Nie Weiran Xu Jun Guo*An Approach to Spam Detection by Naive Bayes Ensemble Based on Decision Induction,* Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on, 16-18 Oct. 2006, Volume: 2, On page(s): 861-866.

[23] Begriche Youcef, Labiod Houda: *A Prior Distribution for Anti-spam Statistical Bayesian Model*. Internationnal Confrerence Network and Service Security: N2S, IFIP & IEEE. Juin 2009, Paris, France.

[24] J.NEVEU: *introduction aux probabilités*. Ecole Polytechnique Département de Mathématiques appliquées