2012 26th IEEE International Conference on Advanced Information Networking and Applications

Evaluating quality of health information sources

Leila Weitzel
Universidade Federal do Pará,
Faculdade de Computação - UFPA
Marabá, Brazil
lmartins@ufpa.br

Paulo Quaresma
Universidade de Évora,
Departamento de Informática
Évora, Portugal
pq@uevora.pt

José Palazzo M. de Oliveira
Universidade Federal do Rio
Grande do Sul - INF/UFRGS
Porto alegre, Brazil

***Abstract*— the Web is an important source for people who are seeking healthcare information. Users, who search for health information online, do so without professional guidance. A major problem faced is the possibility that poor information has detrimental effects on health. In this sense, the main goal of this work is to provide a framework that evaluates health webpage designed especially for common users, those that may lack sufficient knowledge to validate health content. In order to achieve this goal, we proposed a new methodology to calculate a Trust rank based on reputation and a set of quality indicators. The proposed methodology has shown effective to evaluate the quality of health information sources.**

Keywords-metrics information quality; social network analysis; health web-information.

I. INTRODUCTION

The Web is an important source for people who are seeking healthcare information. However, is open to numerous kinds of publishers and information providers. The quality and accuracy is highly variant, highly dynamic, differs in nature, granularity and lifespan. Also contributing to this problem is the fact that, over the last years, the number of users of Online Social Network - OSN like Facebook, LinkedIn, Flickr, MySpace and Twitter has gained increased popularity. This is new space of production, communication, sharing and dissemination of information by giving the opportunity to the users to become collaborating producers (private or public) of the Web content [1].

Users, who search for health information online, do so without professional guidance; besides, they may lack sufficient knowledge and training to evaluate the validity and quality of healthcare web content [2]. It can be worst when, anyone can post information on the Internet regardless of their background, medical qualifications, professional stature, or intention [3], [4]. Many concerns have been raised about the quality of online consumer health information, and the possibility that poor information has detrimental effects on health [5], [6].

The main question is how to trust in web information? Trust is indicative of the confidence placed in a system or entity to deliver desired results [4], [7]. Seen in these terms, trust involves willingness that is not based on having control or power over the other party. Thus, trusting beliefs and intentions reflect the idea of Reputation, and it is only needed when a person is not acquainted with the experts' knowledge [8].

References [4] pointed out that there are many significant factors that affect how users determine trust. Some of them are: *POPULARITY is often correlated with trust but not necessarily; *AUTHORITY can be used by weighting associations and related resources, refers to influences that a user would recognize as proper; because the information therein is thought to be credible and worthy of belief; *REPUTATION comes from direct experience, their own experience or recommendation from others users. The authors argued that, trust can be also influenced by these factors combined together; for example, the association of trusted web site and an unknown source i.e. an article whose author is "John Doe" (distrust authority).

Our first hypothesis is that the process of retweeting creates strong and significant ties between users, based on trust relationship.

Then, our second hypothesis is that Trust, in online health information, can be modeled by a function of < reputation, quality content >. Since reputation is a social evaluation toward a person or a group of people consequently, reputation can be evaluated by Social Network Analysis - SNA. In addition, we also provide a set of indicators for evaluating the quality content. These indicators are grounded in conduct code and technical criteria already established in the literature. It is highlighted that it is not for us to judge medical content.

In this sense, the main goal of this work is to provide a framework that evaluates health webpage designed especially for common users, those that may lack sufficient knowledge to validate or qualify health content.

The following definition is offered for purposes of clarity within this study. *Reputation is the quality of being a "trusted source" or more simply, "credibility". It is used interchangeable with the meaning of the influences (or importance) that a user (or node) comprise; because the content of the information disseminated within her/his social network is thought to be credible and worthy of belief. Our definition of reputation is the amount of trust that a user gives an information source based on previous interactions among them. It must also be stressed that, reputation concept in this paper is not addressed in the context of commercial reputation systems, i.e., between buyers and sellers. Since these systems calculate reputation by a rating "score" which is calculated based on cumulative by its members.

The outline of this paper is as follows: Section 2 presents the background of the research, some related works and we

1550-445X/12 $26.00 © 2012 IEEE
DOI 10.1109/AINA.2012.41

enlighten the problem statement; Section 3 introduces Twitter features and the methodology to measure reputation; In Section 4 we give details of the quality indicators metrics and we also give theses meaning; Section 5 presents the statistical analysis of the data; Section 6 presents others statistical results of nonparametric correlation test. And finally the Section 7 presents the concluding remarks and future work.

II. Background

A. Related works

Several organizations developed quality rating instruments intended to be used by healthcare consumers to evaluate websites. The Agency for Health Care Policy and Research - AHCPR [1] were developed seven criteria to evaluate the quality of health information, which are: *Credibility: includes the source, currency, relevance/utility, and editorial review process for the information; *Content: must be accurate and complete, and an appropriate disclaimer provided; *Disclosure: includes informing the user of the purpose of the site, as well as any profiling or collection of information associated with using the site; *Links: evaluated according to selection, architecture, content, and back linkages; *Design: encompasses accessibility, logical organization (navigability), and internal search capability; *Interactivity: includes feedback mechanisms and means for exchange of information among users, and *Caveats: clarification of whether site function is to market products and services or is a primary information content provider.

The Health on the Net Foundation (HON[2]) created a certification based on a standard conduct code named Net Code of Conduct - HONcode. The code has the intent to allow websites to publish more transparent information. Unfortunately, only a small number of web pages now exhibit the HONcode logo [9]. The principles of HONcode are: *Authoritative: indicate the qualifications of the authors ; *Complementarity: information should support, not replace, the doctor-patient relationship; *Privacy: respect the privacy and confidentiality of personal data submitted to the site by the visitor; *Attribution: cite the source(s) of published information, date medical and health pages; *Justifiability: site must back up claims relating to benefits and performance; *Transparency: accessible presentation, accurate email contact; *Financial: disclosure identify funding sources; and *Advertising policy: clearly distinguish advertising from editorial content.

The Information Access Project[3] - Healthy People 2010 – HP2010 at the United State Department of Health and Human Services identified six properties or types of metadata essential for carrying out a quality evaluation of a health web site, they are: *the identity of owners, developers, and sponsors; *the purpose of the site; *the sources of the content; *the privacy and confidentiality of personal information; *evaluation or feedback mechanisms, and *content update procedures.

Many studies have been conducted to analyze health information on the web throughout the years. Some of them are based on some conduct code cited herein, for example, [10], [11], [2], [12], [13], [13], [1].

B. Problem statement

Web information exists in a large variety of kinds: facts, opinions, stories, interpretations, statistics and is created for many purposes (to inform, to persuade, to sell, to present a viewpoint, and to create or change an attitude or belief). For each of these various kinds and purposes, information exists on many levels of quality and reliability. Different most traditional information media (books, magazines, etc), no one has to approve the content before it is made public. Then, the content can be false or fraudulent, illicit and may contain some printing errors, [1], [3].

The Internet has become a useful education and information tool for healthcare providers and healthcare consumers. Nevertheless, general healthcare consumers must be wary of the legal, quality, and safety implications of relying on the Internet to meet their informational and educational needs. Serious issues must be considered when using the Internet for health and medical information dissemination. Information acquired from Internet have the potential to both improve health and do harm [9], [14].

The studies mentioned herein were performed to establish a methodological framework or a set of criteria on how quality on the web is evaluated in practice, i.e., from the doctor's standpoint. However, trust in information or content is a complex process affected by many factors especially for most ordinary users [4]. There are a few systematic ways available to evaluate the quality of health information, and even less for ordinary user [14].

III. O SN - The Twitter reputation approach

Twitter is an OSN and a web-based Microblogging service that allows registered users to send short status update messages to others. It is a new social software phenomenon that is attracting attention from the popular press [15]. The goal of Twitter is to allow users to communicate and stay connected through the exchange of short messages (up to 140 characters), called "Tweets". According to [16], there is strong evidence that people use them to find information. Twitter provides search interface to easy access public tweets, besides, Bing and Google search engine have both begun to provide online search of Twitter posts.

The Twitters' ties are asymmetric, they are formed basically when a user follow someone, mostly because they are interested in topics that user publishes. The "follower" concept, in Twitter perspective, represents the user who is following you. The "following" concept represents the user who you follow. An interesting feature is when a user posts a tweet, if other users like it, they repost it or "retweet" - RT it; to "retweet" is to repeat/quote someone's tweet. When someone "retweet" you; they are giving you a kind of

[1] http://www.ahrq.gov

[2] http://www.hon.ch

[3] http://phpartners.org/hp/

reputation by sharing your post with their own followers or contacts.

One common type of social analysis is the identification of communities of users with similar interests, and within such communities the identification of the most "influential" users. A simple notion of influence is the number of connections, and influential users act as hubs within their community. The centrality measures of a node importance proposed by [17] are only based on ties (ingoing and outgoing edges) and topological structure of graph. "Edges counts" does not show how important users are. It can be treated only as the "popularity" measure.

Measuring the node importance in a social network has become a worth studying issue. Several works are based on: follower count, co-follower rate (ratio between follower and following), frequency of tweets/updates, who your followers follow, and etc. The centrality measures like degree, betweenness, closeness or eigenvector and either PageRank [18] are proposed in order to tackle with a class of issue. However, sometimes, we must take node importance into full consideration based on several criterions that incorporate more global information. Therefore, evaluating node importance with a single metric can be considered incomplete and limited as it couldn't capture the specific differences among nodes.

In our previous work [19] we proposed a new social network topological structure based on RT weighted ties to rank user influence named Retweet Network or ***RT-network***. We have analyzed the power of retweeting and we also have presented a new methodology to rank nodes based on control weighted parameters. The method was anchored in F-measure to control the weight balance. The experimental results offered an important insight of the relationships among Twitter users. The findings suggested that relations of "friendship" (i.e., users that have reciprocal relationship) are important but not enough to find out how important nodes are. Moreover, centrality measures isolate do not characterize influence but popularity, acting like "edges count".

Thus, in this context, a high-ranking user was characterized by fact that her/his tweets were n-times replayed, the higher it is, the higher Reputation rate.

The equation 1 shows the ranking method proposed by [19]. Let the Reputation be a linear combination of centrality measures with associated weight defined by:

$$\mathrm{Reputation_i} = \frac{\sum_{\mathrm{i}=1}^{m} \ \omega_{\mathrm{i}}}{\sum_{\mathrm{i}=1}^{\mathrm{m}} \ \frac{\omega_{\mathrm{i}}}{x_{\mathrm{i}}}} \qquad (1)$$

$$\left(\sum_{\mathrm{i}=1}^{\mathrm{m}} \omega_{\mathrm{i}}\right) = \ (\delta + \beta + \ \theta + \gamma) = 1 \qquad (2)$$

The ω_i is the weighted parameter. We tested for the values $\delta = 0.25;\ \beta = 0.25;\ \theta = 0.25;\ \gamma = 0.25$.

The parameter $\boldsymbol{x_i}$ corresponds to the centrality measures: Betweenness, Closeness, PageRank [18] and Eigen-Vector, thus $m = 4$.

The RTs posts are marked with characters RT or via @ + "screenname" in the beginning of message, we extracted either both replay tweets and mention.

- *"RT @TheNaturalNews: #Alzheimer's patients treated by playing internet games: http://t.co/dSAmzTv"*
- *"@IRememberBetter: Singing & the Brain: reflections on human capacity 4 music; pilot study of group singing w/ #Alzheimer's http://t.co/0NZXoVU #ArtAlz"*

We extracted the RT post from 152 browsed Twitter's users; in accordance with self Twitter browse interest, in our case we selected health subject. The mining was done during March and April 2011. We crawled [20]about 200 RT per user (this equivalent to about six month of "tweeting") totaling 4350 RT. Reference [21] demonstrated that the median number of **tweets** per user stay between 100 and 1000, emphasizing that maximum tweet values are closely related to the celebrities (actors, singers, pop/rock band, politicians, etc). The authors [21] proved that the majority of users who have fewer than 10 followers never tweeted or did just once and thus the *median stay at 1 tweet per user*. Seen this way, our sample data of RT is perfectly valid. At the end of crawling, we had a ***user-RT database*** of who replayed whom, the relationship between them and the text of retweet. At this point, we could build the RT-network.

The ***RT-network*** was modeled as a direct graph G_{RT} (Figure 1) where each node $u \in V$ (totalling 1237 nodes) represents the users and each edge $a_k = (u_i, u_j) \in A$ represents RT relationship (totalling 1409 edges), i.e., an edge a_k from u_i to u_j stands that user u_i "RETWEET" user u_j .

These edges w_{a_k} between nodes are weighted according the equation 3.

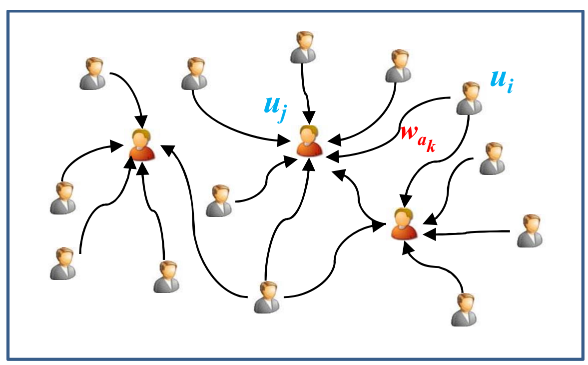

Figure 1. RT-network representing the interactions between users

$$w_{\mathrm{a_k}} = \frac{\sum \mathrm{RT}}{RT_{max}} + \ \alpha \qquad (3)$$

Where $\sum RT$ is the retweet count for $\boldsymbol{u}_{\mathrm{j}}$, and RT_{max} is the maximum number of retweet that user j obtained. The parameter $\boldsymbol{\alpha}$ is a sort of discount rate representing Twitter relationships (follower, following, reciprocally connected and when relationships - follower or following - are absent between users). This parameter is computed according to the ratio of these categories shown in Figure 2.

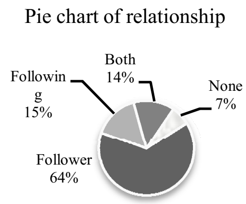

Figure 2. Pie chart of dataset grouping by category

Using this notation, if an individual u_i is a "follower" of u_j, then $\boldsymbol{\alpha} \approx 0.07$ and if is "following" then $\boldsymbol{\alpha} \approx 0.14$, if is both follower and following then $\boldsymbol{\alpha} \approx 0.15$ and if the relationship is absent then $\boldsymbol{\alpha} \approx 0.64$. The parameter $\boldsymbol{\alpha}$ intends to discount the weight of the FOLLOW phenomenon, since many celebrities and mass media have hundreds of thousands of followers.

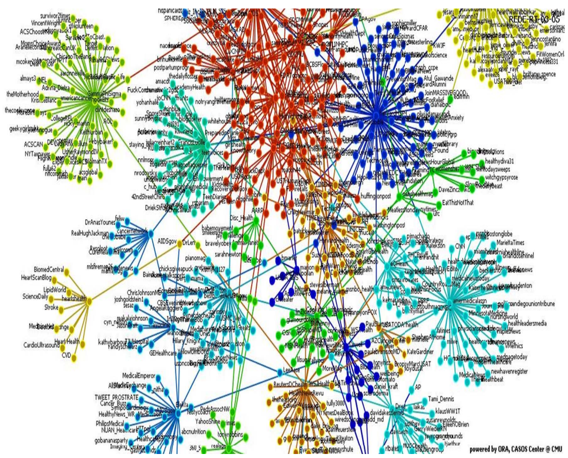

Figure 3. RT-network modeled by Ora[4] Social Network Analysis software

We use Ora Software to compute the measures, and afterward, we compute the ranking methodology proposed by [19] and we archived a ***user influence list*** in descending order by its Reputation (Figure 3).

IV. QUALITY METRICS

Since our research not address medical content, then, the set of indicators proposed is based only on technical criteria and they are medical domain-dependent. We propose ten quality indicators measures QM_j; in what follow we present their definition (and meaning) to avoid misunderstanding:

- INTERACTIVITY (QM_1): if the site give opportunities for feedback and interactivity, i.e. if the web site has email, address, social network association (Twitter, Facebook, MySpace, etc.)
- SEARCH ENGINE (QM_2): internal mechanism to allow user to search content.
- VERIFIABILITY (QM_3): based on philosophical doctrine holding that a statement is meaningful only if it is either empirically verifiable. In that case, if the information can be readily checked.
- SOURCE REFERENCES (QM_4): references or peer-reviewed medical literature, it is imperative to evaluate verifiability; the site should identify the qualifications and credentials of their own and cited authors.
- ADVERTISING POLICY-COMPLETENESS (QM_5): The information is not a substitute for medical advice. A common disclaimer warns users not to use a site to replace traditional health care, representing itself as information rather than a medical-advice source, thus facilitating rather than replacing provider - client interaction.
- HONCODE (QM_6): if the site has HON certification.
- FUNDING AND SPONSORSHIP (QM_7): disclosure of potential conflicts of interest by the site's sponsors. Conflicts of interest may be based on financial dependence, theoretical preference, or intellectual investment, and may indicate bias [5], [12].
- UPDATING (QM_8): Date of Posting, revising of editorial content.
- OTHER SOURCES (QM_9): references to other publications.
- DISCLOSURE OF ADVERTISING POLICY OF PUBLICITY (QM_{10}): publicities can probably carry bias between information and the purpose of the site, for instance, selling medications to promote particular treatments [1], [5].

The criterion QM_j with {j = 1... n} was reported on a binary scale (1-one and 0-zero) defining if an indicator is present or absent in web pages. The indicator QM_8 is continuous-valued indicator, then, it was necessary to re-encode (preprocessing) them in a discrete attribute. There are many ways to realize this process [22]. We choose (for simplification purposes) to employ a discretization with a fixed number of intervals. Therefore, the indicator QM_8 was re-encoded in:

$$QM_8 = \begin{cases} 0.0 \text{ if last update } > 2 \text{ years} \\ 0.5 \text{ if } 1 \text{ year} < \text{last update } \leq 2 \text{ years} \\ 1.0 \text{ if last update } \leq 1 \text{ year} \end{cases}$$

Then, the standing $St(QM_j)$ quality indicator was calculated as:

$$St(QM_j)_i = \left(\frac{\sum_{j=1}^{n} QM_j}{n}\right) \qquad (4)$$

Where $\boldsymbol{j} = 1\ldots\mathbf{n}$ and $\boldsymbol{n} = 10$.

Then, let assume that Trust is denoted by equation 6. The equation 6 is inspired by F-measure. F-measure is generally accepted at Information Retrieval as evaluation performance methods and by far the most widely used. It has been past more than 15 years since the F-measure was first introduced by van Rijsbergen [23]. He states, the F-measure (F) combines Recall (R) and Precision (P) in the following form:

4 http://www.casos.cs.cmu.edu/projects/ora/software.html

$$F(R,P) = \frac{(\beta^2+1)P * R}{\beta^2 P + R} = \frac{1+\beta^2}{\frac{\beta^2}{R}+\frac{1}{P}} \quad \text{where } (0 \leq \beta \leq \infty) \quad (5)$$

Where β is a parameter that controls a balance between P and R. When $\beta = 1$ F comes to equivalent to the harmonic mean of P and R. If $\beta > 1$, F becomes more recall-oriented and if $\beta < 1$, it becomes more precision oriented $F_0 = P$. The possible scores for this rank range from 0.0 to 1.0. A score of 1.0 would be the highest position, and a score of 0.0 would indicate lowest position.

$$\text{Trust} = \left(\frac{2}{\frac{1}{St(QM_j)} + \frac{1}{Reputation}} \right) \quad (6)$$

Substituting (1) and (4) into (6) then, equation (7) yields:

$$\text{Trust} = \left(\frac{2}{\frac{1}{\left(\frac{\sum_{j=1}^{n} QM_j}{n}\right)} + \frac{1}{\left(\frac{\sum_{i=1}^{m} \omega_i}{\sum_{i=1}^{m} \frac{\omega_i}{x_i}}\right)}} \right) \quad (7)$$

We were motivated to formulate the Trust by using harmonic means notion. Mostly because, simple mean are sensitive to outliers thus, sometimes means does not reflects the quantity desired [24].

V. Statistical Analysis

We systematically have examined the ***user influence list***. We found that 70 percent have site (or blog) of human health. In this sense, we selected only those users and discard the remainders; afterward, we sorted them again by their Reputation. And the top 21 list users are shown in Table 1 column three the ***user influence list n.2***

The following statistics were calculated only for the Top 21 users. We found some interesting features:

Approximately 10 percent of site or blog did not have your own content, i.e. when users search for a particular subject, the search engine brings them out of the original site.

The majority of URL is a sub-domain of CDC.gov URL that archived the first position in Table I. Except for the lines: 16, 18, 19 and 20 in Table 1 all of the remainder (81percent) are Government website or blogs.

About 5 percent have HONcode. It is interesting to notice that, MQ_6 is inversely proportional to the category of site, i.e., although 81 percent are government or public websites, they do not have the certification logo of HONcode.

About 14 percent have Private Funding or Sponsorship (QM_7) and just 19 percent have Complementary Sources (QM_9). It must be stressed that, we found high positive correlation (0.78) between Disclosure of Advertising Policy Publicity (QM_{10}) and Funding/Sponsorship (QM_7). This indicates that may have conflict of interest and the financial relationships have potential to bias the information distributed. We also noticed that there is a negative correlation (-0.68) between HONcode (QM_6) and Search Engine (QM_2). We compute the acceptance rate of retweet ϕ (RT), which is calculated using the following equation 8:

$$\phi\,(\text{RT})_i = \left(\frac{\text{RT}}{\text{T}}\right) \quad (8)$$

Where RT is the total of retweet, the T is the total of tweet of user ***i***. The lowest value was 0.19 percent and the highest value was 31 percent (mean equal to 7 percent).

VI. Assessing Trust

For all users in the ***user influence list n.2*** we calculate the quality indicators as proposed by equation 4, and the results are shown in Table 1 column four.

Finally, we compute de ***trust list*** based on equation 6 (shown in Table 1, column five). The results are classified descending order of Trust parameter in Table I. Rather than evaluating the values calculated directly, we sorted the URL by each list, so that the rank of 1 indicates the most influential user and the twenty one positions is less influential. Then, we compute the Kendall Tau ($\boldsymbol{\tau}$) Correlation and Spearman-Rho Rank test (Rho =$\boldsymbol{\rho}$) for the two listing Reputation, Trust and $St(QM_k)_i$ rank. The correlations rates are shown in Table II.

TABLE I. Results classified by Trust rank descending order

Pos.	**URL**	***Reputation***	***St(QMk)i***	***Trust***
1	cdc.gov	1,00	0,75	0,85
2	cdc.gov/socialmedia	0,96	0,70	0,81
3	womenshealth.gov	0,80	0,80	0,80
4	blog.aids.gov	0,90	0,70	0,78
5	cdc.gov/cancer	0,83	0,70	0,76
6	ndep.nih.gov	0,70	0,80	0,74
7	emergency.cdc.gov	0,73	0,70	0,71
8	hhs.gov	0,86	0,60	0,70
9	girlshealth.gov	0,76	0,60	0,67
10	cancer.org	0,53	0,90	0,67
11	flu.gov	0,56	0,80	0,66
12	cdc.gov/niosh	0,60	0,70	0,64
13	cdcnpin.org	0,50	0,90	0,64
14	cdc.gov/ncbddd/actearly	0,30	0,70	0,42
15	nineandahalfminutes.org	0,66	0,30	0,41
16	heart.org	0,26	0,60	0,36
17	cdc.gov/flu	0,23	0,70	0,35
18	drweilblog.com	0,20	0,90	0,32
19	health.com/health	0,33	0,30	0,31
20	health.discovery.com	0,46	0,20	0,28
21	healthcare.gov	0,16	0,50	0,25

The Kendall Tau ($\boldsymbol{\tau}$) Correlation and Spearman-Rho Rank test (Rho =$\boldsymbol{\rho}$ **-** Equation 9) are the two most commonly used nonparametric measures of association for two random variables [25]. The two tests are similar; they compute a correlation coefficient by ranking all possible pairs of entries.

Kendall's Tau is based on scores assigned to each pair of bivariate observations, say (x_1, y_1), (x_2, y_2) that measure the concordance between the two observations. In short, (x_1, y_1) and (x_2, y_2) are said to be concordant if: $x_1 > x_2$ and $y1 > y2$ or if $x_1 < x_2$ and $y_1 < y_2$. They are discordant if: $x_1 > x_2$ and $y_1 < y_2$ or if $x_1 < x_2$ and $y_1 > y_2$. Concordant pairs are assigned a score of 1, discordant pairs are assigned a score of -1, and pairs in which there is equality among either variable are assigned a score of 0.

The Spearman correlation is the ordinary (Pearson) correlation coefficient of the transformed random variables F(X) and G(Y). It assesses how well the relationship between two variables can be described using a monotonic function. The coefficient is also inside interval [-1, 1].

$$\boldsymbol{\rho} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \tag{9}$$

The null hypothesis H_0 tested states that "there is no difference between Reputation and $St(QM_k)_i$ rank".

TABLE II. CORRELATION RESULTS BETWEEN *$St(QM_k)i$ AND RANK*

		$St(QM_k)i$	***Reputation***
Pearson	***$St(QM_k)_i$***	1.00	0.816
	Reputation	0.816	1.00
Kendall's Tau	***$St(QM_k)_i$***	1.00	0.820
	Reputation	0.820	1.00
Spearman's rho	***$St(QM_k)_i$***	1.00	0.894
	Reputation	0.894	1.00
All Correlation is significant at the 0.01 level (1-tailed)			

As we expected, there are a strong positive correlation between Reputation and $St(QM_k)_i$ the values range between 0.816 and 0.894. Then, the null hypothesis cannot be rejected at significance level of 99 percent. We computed the canonical correlation - CC to assess the relationship between these variables. The results suggest that the relationship between Reputation and $St(QM_k)_i$ is CC = 0.8156, p = 0.000 percent is statistically significant.

Alexa[5] is website that routinely (more specifically daily) computes the Alexa Traffic Rank – ATR, which is a ranking of the most visited websites. The rank is calculated using a combination of average daily visitors and pageviews over the past three months. ATR is also a measure of American Medical Association [6] popularity. Alexa categorizes all websites into three sets: global, by country and by category. On May 31 (2011), we extracted the ATR rank list by health category, then, we compute the correlation between Trust and ATR, and the results are shown in Table 3.

[5] http://www.alexa.com/

[6] http://www.ama-assn.org/

TABLE III. CORRELATION RESULTS BETWEEN ALEXA RANK *AND REPUTATION RANK*

		Alexa	***Trust***
Pearson	***Alexa***	1.00	0.096
	Trust	0.096	1.00
Kendall's Tau	***Alexa***	1.00	0.13
	Trust	0.13	1.00
Spearman's rho	***Alexa***	1.00	0.096
	Trust	0.096	1.00
All Correlation is significant at the 0.01 level (1-tailed)			

The null hypothesis H_0 states that "there is no difference between ATR rank and Trust rank". With 0.096 percent there is no difference between these ranks, as a result we then reject the null hypothesis at significance level of 99 percent.

The Table IV shows the correlation matrix of the studied parameters. The Trust factor proposed is strongly correlated to Betweenness Centrality measure (81 percent), and weakly correlated to Closeness Centrality (zero percent). It must be highlighted that Alexa rank are weakly correlated to Betweenness, Closeness, PageRank or Eigen-Vector, mostly because they rank nodes using a combination of average daily visitors and pageviews over the past three months nodes by .

TABLE IV. CORRELATION MATRIX

%	**Betweenness**	**Closeness**	**Eigen-vector**	**Page Rank**	**Trust**	**Rank (ATR)**
Betweenness	100	-2	29	36	**81**	0
Trust	81	0	40	40	100	-6
Closeness	-2	100	11	14	**0**	-10
Eigen-Vector	29	11	100	92	40	-34
Page Rank	36	14	92	100	40	-46
Reputation	-83	-9	-50	-51	-97	10
Rank (ATR)	0	-10	-34	-46	-57	100

VII. CONCLUDING REMARKS AND FUTURE WORK

The proposed methodology is based on credible and worthy of belief source, and we named it Trust. Our hypothesis was that Trust can be modeled by a pair reputation factor and a set of quality indicators.

We utilized the SNA to figure out user's reputation, since SN creates trust between agents because they allow their members to learn about each other through repeated interactions. In that case, the interaction can be done by "retweeting" process. Finally, we proposed a set of technical indicators for evaluate the health quality content.

Thus, the major contribution of this work was mostly providing a framework to evaluate health webpage in particular for ordinary user.

Some interesting findings were figured out: the great majority health information available is from government or public sources, and these sites do not have the logo of HONcode.

We noticed strong and positive correlation between Disclosure of Advertising Policy Publicity (QM_{10}) and Funding/Sponsorship (QM_7). This means that it depends, in part on how well conflict of interest is handled, in the case that the site has financial or relationships that inappropriately influence (bias) their actions (also known as competing interests). If sites receive funding from commercial firms or private foundations, then the conditions of this funding have the potential to bias and otherwise discredit the information presented. According to reference [13] websites do not provide enough information for visitors to assess whether a conflict of interest with pharmaceutical companies exists.

The negative correlation between HONcode (QM_6) and Search Engine (QM_2), seems to be a perfectly reasonable, since their own content must be evaluated in order to achieve the logo certification.

The computed acceptance rate ϕ (RT) showed that the "retweeting" process was not broadly adopted. This may mean that users are more discerning when choosing what or who to retweet; in spite of the percentage of follower relationship reached about 64 percent the average of retweet was less than a one unit per user. Indeed, these finding reinforce our first hypothesis.

On the internet, we have to use a different mechanism to decide what sources of information are trustworthy – everyone is, or could be an authority or expert. In this sense, from the information consumer's perspective, a very reasonable reversal of the real life process is taking place, the suggestion is that we do not necessarily trust a source of information (e.g. a "*John Doe* Expert Blog") just because it exists, but we increase our level of trust as we realize that a consistent quality of information is being delivered.

Our main goal was to provide a framework that evaluates health webpage designed especially for ordinary users those that may lack sufficient knowledge to validate or qualify health content. In order to achieve this goal, we have formulated two hypotheses: first is that retweet creates strong and important ties between users in Twitter microblogging, and second we can model Trust by a function of Reputation factor and a set of quality indicators. The Reputation issue was rooted in SNA and the quality indicators were based on code conduct and technical criteria.

Taking into account all the statistical results and inferences we can state that our methodology to modeling trust is suitable.

Acknowledgment

Leila Weitzel and José Palazzo M. de Oliveira would thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brazilian National Research Council) for the partial financial support of this work.31/2010-9. Leila Weitzel would also like to thank Universidade Federal do Pará for providing partial financial support of this work.

References

[1] B. Stvilia, L. Mon, e Y. J. Yi, "A model for online consumer health information quality", *Journal of the American Society for Information Science and Technology*, vol. 60, nº. 9, p. 1781-1791, 2009.

[2] G. Eysenbach, "How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews", *BMJ*, vol. 324, p. 573-577, mar. 2002.

[3] J. Bar-Ilan e B. C. Peritz, "A method for measuring the evolution of a topic on the Web: The case of 'informetrics'", *Journal of the American Society for Information Science and Technology*, vol. 60, nº. 9, p. 1730-1740, 2009.

[4] Y. Gil e D. Artz, "Towards Content Trust of Web Resources", *Journal of Web Semantics: Preprint Server*, vol. 5, nº. 4, 2011.

[5] J. A. Powell, M. Darvell, e J. A. M. Gray, "The doctor, the patient and the world-wide web: how the internet is changing healthcare", *J R Soc Med*, vol. 96, nº. 2, p. 74-76, 2003.

[6] G. Eysenbach et al., "Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information.", *BMJ*, vol. 317, nº. 7171, p. 1496-1502, 1998.

[7] J. Golbeck, B. Parsia, e J. Hendler, "Trust Networks on the Semantic Web", in *In Proceedings of Cooperative Intelligent Agents*, 2003, p. 238–249.

[8] J. H. Numan, "Knowledge-Based Systems as CompanionsTrust, Human Computer Interaction and Complex Systems", University of Groningen, Netherlands, 1998.

[9] Y. Wang e Z. Liu, "Automatic detecting indicators for quality of health information on the Web", *International Journal of Medical Informatics*, vol. 76, nº. 8, p. 575-582, ago. 2007.

[10] G. Eysenbach, "Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web: A Systematic Review", *JAMA: The Journal of the American Medical Association*, vol. 287, p. 2691-2700, maio 2002.

[11] A. Risk e J. Dzenowagis, "Review Of Internet Health Information Quality Initiatives", *Journal of Medical Internet Research*, vol. 3, p. e28, dez. 2001.

[12] J. G. Anderson, "Consumers of e-Health: Patterns of Use and Barriers", *Social Science Computer Review*, vol. 22, p. 242-248, maio 2004.

[13] D. Ball, K. Tisocki, e A. Herxheimer, "Advertising and disclosure of funding on patient organisation websites: a cross-sectional survey", *BMC Public Health*, vol. 6, nº. 1, p. 201, 2006.

[14] H. W. W. Potts e J. C. Wyatt, "Survey of Doctors' Experience of Patients Using the Internet", *Journal of Medical Internet Research*, vol. 4, nº. 1, p. e5, mar. 2002.

[15] C. Honey e S. C. Herring, "Beyond Microblogging: Conversation and Collaboration via Twitter", in *System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on*, 2009, p. 1-10.

[16] J. Teevan, D. Ramage, e M. R. Morris, "#TwitterSearch: a comparison of microblog search and web search", in *Proceedings of the fourth ACM international conference on Web search and data mining*, Hong Kong, China, 2011, p. 35-44.

[17] R. Albert e A. L. Barabási, "Statistical mechanics of complex networks", *Reviews of modern physics*, vol. 74, nº. 1, p. 47, 2002.

[18] L. Page, S. Brin, R. Motwani, e T. Winograd, "The PageRank Citation Ranking: Bringing Order to the Web.", Stanford InfoLab, Technical Report, 1999.

[19] L. Weitzel, P. Quaresma, e J. O. Palazzo, "MEASURING NODE IMPORTANCE: A MULTI-CRITERIA APPROACH", in *WWW/INTERNET 2011*, Rio de Janeiro, 2011.

[20] A. Jhoseph da Silva de Araujo e W. Fernando Garcia Rodriguez, *JCrawler4T*. Pará: Universidade Federal do Pará, 2011.

[21] A. Pal e S. Counts, “Identifying topical authorities in microblogs”, in *Proceedings of the fourth ACM international conference on Web search and data mining*, Hong Kong, China, 2011, p. 45-54.

[22] H. Liu, F. Hussain, C. L. Tan, e M. Dash, “Discretization: An Enabling Technique”, *Data Mining and Knowledge Discovery*, vol. 6, nº. 4, p. 393-423, out. 2002.

[23] C. J. van Rijsbergen, *Information retrieval*, 2º ed. London: Butterworths, 1979.

[24] P. Bullen, *Handbook of means and their inequalities*. Dordrecht ;;Boston: Kluwer Academic Publishers, 2003.

[25] C. Genest e J. Neslehov, “Analytical proofs of classical inequalities between Spearman’s [rho] and Kendall’s [tau]”, *Journal of Statistical Planning and Inference*, vol. 139, nº. 11, p. 3795-3798, nov. 2009.