uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Hicham Ibrahim

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Electrical Engineering)

GRADE / DEGREE

School of Information Technology and Engineering

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Semantic Similarity based Trust Computation in Websites

TITRE DE LA THÈSE / TITLE OF THESIS

Prof. A. El Saddik

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Prof. D. Ionescu

Prof. M. Frize

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Semantic Similarity based Trust Computation in Websites

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.Sc. degree in
Electrical and Computer Engineering

School of Information Technology and Engineering
Faculty of Engineering
University of Ottawa

Hicham Ibrahim, Ottawa, Canada, 2008

Your file Votre référence
ISBN: 978-0-494-48465-4
Our file Notre référence
ISBN: 978-0-494-48465-4

Abstract

The World Wide Web has evolved at an extreme rate due to its capacity to provide an endless amount of information to the public users. Hence, users find themselves lost in a pool of information, without knowing what to trust in order to retain. They normally access this information from the first few links provided to them by the search engines such as Google or Yahoo. However, this factual information, which is essentially popular due to higher hit counts, may not always be trustworthy from the factual aspects. In other words, we can ask ourselves if one can actually trust specific information by the simple fact that it has been ranked high by virtue of being accessed frequently by other users.

This thesis presents a mechanism to determine the trust of websites, even though they are not so popular, based on the semantic similarity of their content with the already established and trusted websites. The proposed method allows the dynamic computation of the trust level of websites, and hence overcomes the dependency on the traditional user feedback methods for determining the trust. In fact, our method attempts to emulate the evolving process of trust that takes place in a user's mind. The experimental results have been provided to demonstrate the utility and practicality of the proposed method.

Acknowledgements

This thesis is the result of two years of work during which I have been accompanied and supported by several people, who helped me and guided throughout the entire process. It is now with great pleasure that I this opportunity to sincerely thank them.

Firstly, I would like thank the University of Ottawa for giving me the wonderful opportunity and the financial means to pursue my studies and obtain a Master's degree in Computer and Electrical Engineering.

My most earnest and deepest acknowledgments must go to my supervisor, Prof Abdulmotaleb El-Saddik, and my co-supervisor Prof Pradeep Kumar Atrey, for their constant and admirable motivation, help, and support. The completion of this work would not have been a success without their unquestionable belief in me, and without their immeasurable patience. Their role in this thesis was without a doubt as instrumental as mine. Thank you again for believing in me and never leaving me discouraged! Moreover, I would like to thank my colleagues at the MCRLab for their mental and emotional support, not to mention their much appreciated feedback.

Finally, I would also like to thank the ten users (whose names will be omitted for confidentiality purposes) who participated in our Empirical User Study for a period of three months out of their own will and personal interest in the topic. Thank you very much for your time, patience, and collaboration - it was tremendously appreciated! The study would not have been the same without your unforgettable collaboration. Thank you again!

Contents

List of Figures

List of Tables

List of Symbols

W_1, W_2	*Websites* – website 1, website 2.
$\phi^k_{ij}(t)$	*Semantic Similarity Coefficient or Similarity Measure* between two websites W_i and W_j for a domain k at the time instant t.
$LSA_{ij}{}^k(t)$	*Current Similarity Value* between W_i and W_j for a domain k at the time instant t.
$T_i(t)$	The overall trust $T_i(t)$ for a particular website W_i at a time instant t.
$T^k_j(t)$	*Trust level* of website W_j within a particular domain k at time t.
T_{spec}.	*Threshold* for trust level of any trusted website.
$\beta(t)$	*Growth Factor* for the trust level T^k_i of website W_i for domain k.
$\Delta\phi^k_{ij}(t)$	*Difference or change in the Similarity Measures* between website W_i and website W_j for domain k at time t.

μ, r_{growth}	*Growth Value* used to control the rate of growth or decay, depending if similarity measure, $0 < \Delta\phi^k_{ij}(t) > 0$.
$N(t)$	*Normalization factor* used to limit the trust level value between [0,1].
D^k_i	*Mean Square Difference* (or *Individual Difference*) between the trust level of a website W_i for domain k when our method and the traditional user feedback method are used.
D_i	*Average Difference* in trust levels between our method and user feedback method for one domain within a website.
D_{all}	*Overall Average Difference* in trust levels between our method and user feedback method.

Chapter 1

Introduction

1.1 Motivation

The past decade has seen extraordinary growth of the World Wide Web. It is mainly because of its ability to provide an enormous amount of information to the users. In general, web users can find information of their interest from several websites. For example, if someone looks for the news belonging to politics, there could be several websites such as www.cnn.com, www.bbc.com, www.cbc.ca, etc. These websites also provide general news in several other domains such as health, sports, entertainment, business, etc. Furthermore, one can visit specified websites that concentrate in particular domains such as health (e.g. familDoctor.org, kidsHealth.org, HealthFinder.gov, etc.) or sports (e.g. tsn.ca, espn.go.com, skySports.com, etc.).

The availability of a large number of websites belonging to a particular domain raises the issue of which website should be trusted, since not all the websites always provide factual and trustworthy information. For example, for political news, the website www.cbc.ca may be trustworthy, but any arbitrary website may not be trustworthy, or vice versa. Moreover, the existing search engines and technologies (such as Google) do not perform trust-based searching, but they tend to reward highly popular sites by listing them more than the less popular sites. The popularity of a website is usually determined based on various criteria such as hit count, etc. [1], that does not truly reflect its trust level. For instance, a website with a very low hit count will have a low trust level despite the fact that its actual content may be quite similar to a trustworthy

website. Therefore, the fact that a website has a high hit count does not necessarily imply that it is trusted among people. This website could be seen as a "popular" one, but the popularity does not necessarily always imply trustworthiness.

The rewarding strategy such as "hit count" adopted by the existing search mechanisms makes it quite difficult to prevent these websites from dominating the global exchange of ideas on the Internet. Due to this, several important trustable websites go unnoticed and the users tend to limit their exploration to a few websites such as MySpace, Yahoo, eBay, for instance. This could definitely raise more issues with search engines that seem to amplify this tendency by ranking search results partially based on a particular site's current popularity. Thus, highly popular sites could become even more popular, whereas less popular sites could fade even more in the World Wide Web. This could lead one to detect a bias or inequitable element introduced by the more popular search engines which seem to discriminate against these so called smaller and less popular, yet trustable, websites.

This inequality could be fought by simple human curiosity: a more experienced user enters well-thought specific search terms that would narrow down the results and redirect him to smaller and less popular sites giving him the specific information that he needs; information that a highly popular site could not necessarily provide. This would be ideal if most users were skilled in web searching, and knew the right combination of words to make a more efficient and optimal search that would lead them to the less popular but trustable sites. However, the majority of these users are novice: they either lack these "not-easily acquired" searching techniques that need to be taught in some cases by more experienced users, or simply lack this human curiosity that could stop the biasness by search engines. In other words, most people would rather search, explore and investigate less, by sticking to what they know best, the popular internet sources, when these sources could be in fact untrustworthy. In fact, our method attempts to eliminate these discriminatory elements, and makes it easier for the somewhat indolent users to obtain credible and trustworthy information.

1.2 Trust: various perspectives

We can start this section by asking ourselves this question: how do we know if we should trust the information presented in a site from the factual aspect? Before addressing this issue, we will try to define what we mean by "trust" when it comes to internet page content. In fact, the next section provides several related works that present numerous definitions of internet trust, given from different authors and different disciplines. However, most of these definitions, just like the majority of the related works presented, are somewhat tailored to reflect online trust for E-commerce systems rather than the factual information presented on a webpage. So when we usually hear the word trust and trustworthiness on the web, there is some sort of financial transaction involved. Yet, these definitions give us a general guideline of what "trust" really means, and could be adapted to not only reflect online trust with regards to E-commerce transactions, but to reflect online trust with regards to factual information on the Web. This being said, we can simply define trust as a measure of interaction by the user, or the measure of the willingness of a user to interact with a web page, believe the information presented in that page, and act on it [16]. Now this "willingness" could be somewhat difficult to assess or calculate. However, if we base this assessment on empirical data combined with an efficient mathematical model, we can indeed attribute a numerical value to this trust.

How do we trust the factual information presented in a certain informative web page? There are several general guidelines that we can adhere to, but they do not provide us a quantitative measure of the trust. These could be used to influence the trust one might have in a website's informational content. First, we look at several higher level elements that could be more difficult to identify or assess in some ways:

- Other pages or links viewed on that site.
- Other sites that one regards as related and present similar or comparable relative information.
- Any typical information one might know about the organization owning the site.
- The credibility of the owner of the site.

- The credibility of the top level domain of the site (e.g. .gov, .edu, .org, etc.) or its external certification.
- The origin of the content provided by the author.
- The intention of the author or of the source of information.
- The credibility and expertise of the author relative to the information presented.
- A clear differentiation between opinions and facts.
- The objectiveness of the information presented.

Other more specific lower level, and somewhat less subtle, elements include:

- The availability of the website – whether it's running or not.
- The grammatical content.
- The amount of broken links.
- The name of the author (i.e. makes the page less impersonal).
- The providing of a "Contact page", an "About" page, or a "FAQ" page.

As mentioned, the latter elements or guidelines could present a certain difficulty in their consideration or assessment. For this reason, one can try to find certain credible and objective reviews made by a particular expert in the domain, or by respectable institutes, that recommend particular websites that have trustworthy information. For instance, the Medical Library Association website (MLA) [2] proposes ten trustworthy websites that provide health information with regards to several health issues. These include {cancer, healthFinder, MedlinePlus, cdc}.gov, {familyDoctor, kidsHealth, noah-health}.org, HIVinSite.edu, {mayoClinic, MEDEM}.com [2]. This website also gives a link to the Consumer and Patient Health Information Section (CAPHIS) [17] of MLA which evaluates web sites based on the following criteria: credibility, sponsorship/authorship, content, audience, currency, disclosure, purpose, links, design, interactivity, and disclaimers [17]. But once again, the user has to be vigilant in his findings and personal evaluation of such sites. Also, the sites that make such recommendations and evaluations are quite rare to find, especially when it comes to information about topics somewhat less vital than health, but still have a certain degree of importance to the general public.

All this leads us to believe that trust is constructed by the site owner which has to be a member of something, which in turn is a member of something, which in turn is a member of something that the visitor has heard of, and can trust. This evolution or built of trust could be seen as chain reaction of "favorable associations" [16]. As mentioned previously, in order to properly provide a quantitative measure to these associations, and thus attribute a certain trust level to the factual content of a webpage, one has to include empirical data (e.g. user surveys, and user feedback method), which subconsciously examines the first 'higher level elements' presented earlier, and incorporate this data to a certain computational model that attempts to depict this trust as close as it can to human judgment. And, this is indeed what we have tried to achieve in this work – dynamically evolve the trust level in a particular website based on how similar it is found to be compared to a trusted website in the same domain. As a matter of fact, this 'comparison to a trusted website', being an indispensable component of our proposed method, is the second and probably the most important element in the 'higher level elements' list presented above. Essentially, our method would attempt to emulate, as much as possible, how this evolving process of trust would progress in an actual user's mind.

It is important to note once again that this trust level is different from the trust computed based on privacy and security, as the former implies the user's trust in the factual aspect of the information and the latter is mainly based on its cryptographic security. Note that we define this factual information as being any information indicating events or things that are known to have happened or existed, or any provable truth. Therefore, we refer to the factual trust level as the trust level.

1.3 Importance of Computing Trust

Determining the trust level of the websites would be of significant help in performing trust-based searches. Practically, trust levels could further advance in several information technologies such as search engines, text corpus visualizations, and a variety of applications that deal with information filtering, sorting and retrieval, all being information that rely greatly on similarity measures. In general, this trust level can be practical for text handling, web semantics, data

processing, and data mining. Furthermore, this would allow one to have trust in internet domains such as health, where the accuracy of the medical information becomes indispensable. This is a domain where inaccurate or erroneous facts could obviously compromise one's quality of life and consequently lead to severe repercussions in the extreme scenarios. Finally, the computed trust level is important for students or anyone in general with uncertainties regarding the finding and the use of certain references (e.g. for a school project). This could be a way to provide a credible and less familiar source of information the trust level that it deserves.

Also, having a trust level for the websites is very important as it allows web users to see the web information along with its trust level. The website can have different trust levels for different domains or for a specific domain within a certain domain. For example, in general news domain, cnn.com may be highly trusted in the political news domain, but may be less trusted in the entertainment news domain. Or, in the health domain for instance, familyDoctor.org may be highly trusted in general information concerning diseases such as diabetes and cardiovascular diseases, but less trusted in specifics about more complex diseases such as cancer and HIV.

1.4 Thesis Contributions

In this work, we propose a method to dynamically compute the trust level of a website of a particular domain based on how similar its content is with a trusted website of the same domain. This approach could be illustrated by a simple example: let us assume that one consults a certain website for the sports news, and thus builds a certain confidence or trust in this website over a certain period of time. Then, upon consulting another website for the sport news simultaneously and observing the similarity/dissimilarity between the two, one will eventually develop a certain trust in the second website if the content of both websites are found to be similar over this time frame. Thus, the computation of this trust level will be based on the past history of similarity measured between this web site and a so called "trusted" website in that domain. The similarity measured between the two websites is computed based on how similar the content of the two websites has been in the past [29]. The so called "trusted" website within a particular domain could be found by using traditional methods such as hit counts, user surveys, etc.

The main contribution of this thesis can be summarized as follows: we propose a method to dynamically compute the trust level of a website based on the history of semantic similarity of its content with the other already well-trusted website. The proposed method determines the trust that is based on factual aspects of the website. This can help in improving the performance of systems involving decision support, searching (which includes filtering, sorting, indexing, etc..), or any other text based systems. It can also serve as a way to identify credible references for information use in general (e.g. report, paper, or journal writing).

1.5 Thesis Organization

This thesis is organized as follows. In Chapter 2, we present a review of the previous and current works related to 'online trust'. The somewhat ambiguous notion of 'trust' in these works mostly involves E-commerce sites, where financial transactions are required. However, while discussing this, the general notion of trust, and how it's perceived within different disciplines, will also be covered. This allows us to acquire an overall understanding of trust, and then attempt to relate this to our work specifically, as several of the trust-related factors induced from these different disciplines are indeed relevant to our proposed work in this thesis. Furthermore, several works related to 'reputation' and 'confidence' will also be covered, considering the direct association between these terms and the notion of trust.

Chapter 3 presents the proposed method which determines the trust in a new non-trusted website using its "Similarity Measure" with the trusted website(s). The similarity measure between the two websites is computed based on a model named Latent Semantic Analysis (LSA), which will also be covered in this chapter. Furthermore, an examination of the overall system architecture, its integral components, and the way these components communicate, will be given.

In order to demonstrate the utility of the proposed method, the experimental results are presented in Chapter 4. This chapter begins with a description of the data set user survey performed in order to determine the trusted website. Then, the Empirical User Study or User Feedback Method, and the rationale behind it, is clearly described. Finally, several analyses are made with

the results from our method. In particular, the results of our Empirical User Study are critically compared to those of our proposed method, in order to illustrate the strength of the latter. Also, several aspects limiting the performance of our method are addressed in this chapter.

Chapter 5 presents summary and conclusions of this dissertation. Moreover, certain limitations to affecting our method's performance are addressed, followed by potential solutions in order to enhance this performance. Lastly, future work related to the topic is discussed as a final point of this thesis.

Chapter 2

Literature Review

In order to properly develop this section, we will consider three key words - trust, reputation, and confidence, which are related to each other but have been used in different computing contexts. The definition of trust includes the word confidence, and vice versa [18]. Hence, the two words can be used interchangeably. On the other hand, reputation can be seen as an estimation used to build trust or confidence. Although similar and related in definition, the research communities have used them differently in different computing contexts. This difference will allow us to describe related works that are associated with these words and which could be directly or indirectly related to the problem addressed in this thesis.

It is important to note that there are several works related to these key words or concepts considering the wideness of these notions. Covering every work, weather directly related or indirectly related to this thesis topic, would be quite extensive. For that reason, we will only present in greater details the works that we consider being of higher relevancy than others to our topic, and briefly examine the other ones. This being said, concentrating on previous works that are relevant to the concept of "trust", instead of reputation or confidence for example, would only be logical considering the "something" between this notion and the topic of this thesis, and the importance to fully grasp the meaning of the concept, one that was shown to be somewhat ambiguous.

2.1 Online Trust

The word trust is defined by the Oxford English Dictionary as confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement [18]. Furthermore, an analysis of the term by Nissenbaum[20] led to this following statement "Trust is an extraordinary rich concept, covering a variety of relationships, conjoining a variety of objects. One can trust (or distrust) persons, institutions, governments, information, deities, physical things, systems, and more" [20]. Even basic definitions of the word have caused several disagreements in the research world [19]. In fact, we can attribute this literal divergence among scholars to the multiple definitions that the word presents. On one hand, trust can be viewed as an abstract notion, where its use can be interchanged with words such as confidence, reliability or credibility. Therefore, simply making a clear and comprehensible distinction between trust and its related notions shows to be quite an obstacle for researchers across the world [5]. On the other hand, trust can be viewed as "...a multi-faced concept that incorporates cognitive, emotional, and behavioral dimensions" [5]. Several studies emerging from different disciplines have been made regarding the concept of trust, while respectively presenting a distinctive understanding of the concept.

In a very useful guide on the trustworthiness of websites from the Google Librarian Center, Karen G. Schneider [3], Director of Librarians' Internet Index (LII), presents a five-point system to determine which website one can trust when a search engine displays hundreds of websites corresponding to the desired topic. This five-point system includes the following principals: availability, credibility, authorship, external links, and legality. Here's a description of each principle from Ms. Schneider's work:

- Availability: The first question presented within this principle is quite simple and self explanatory "Is the site up and running?"[3]. The second question, somewhat more complex and ambiguous, is relative to whether or not the information presented on a website is free. The main idea here is that, perhaps one should question a website if there is a certain registration fee in order to access specific information that is hidden "behind walls of one sort or another"[3]. Also, this required registration raises uncertainty on the use of personal information by the website. An advice given by the author is to browse through "...the website's mission statement, 'About' page, or registration sign-up page..." to establish whether or not this hidden information is trustable and worth one's time and money.

- Credibility: In this principle, we can ask the following question "Does the website contribute current, accurate information? Is the site author(s) qualified to present the content provided?" [3]. For their Librarian's Internet Index (LII), Ms. Schneider and her colleagues discarded several websites based on their lack of credibility based on a credential insufficiency by the author with regards to the information presented or simply based on a lower-grade content ("...recipes that misstated quantities, or definitions we knew to be wrong"[3]). She also warns about personal websites, who need to be examined thoroughly before any facts or suggestions that they may offer is taken into account. Hence, it is always important to "look for an About page or an author biography" [3]. Furthermore, Ms. Schneider suggests in some cases to limit searches "...to .edu,.gov, .org, and .ca domains to quickly winnow the search to sites known to be authoritative." [3].

- Authorship: Ms. Schneider and her team at the LII express their skepticism about grammar and typographical errors found in websites, and how these errors are clear indications of doubtful content. An exception to this rule of thumb includes websites translated from languages other than English. In turn, this exception is valid only if one can verify and confirm that the original language that the website was written in is free of typos and grammatical errors. An advice would be to test a few websites through a spell-check in any word processing program if uncertainty in the content arises.

- External Links: The issue put forth in this principle is one of broken links, whether leading to different sites than indicated or simply being inactive. The guide points out how these broken links (i.e. many broken links, not just one or two) significantly decrease the website's reputation, and how they represent a lack of awareness by the author with respect to the presented content in the actual webpage or website. To ensure maintenance of links by a certain website, Ms. Schneider suggests looking carefully for indicators such as dates showing the last time a particular webpage was updated. Finally, the librarian advises to "...beware of student project web sites, and personal web pages with many, many links!" [3].

- Legality: This principle examines the legitimacy of a certain website in ensuring that its author possesses the legal rights to publish certain information while respecting "...copyrights and fair use guidelines."[3]. The guide gives a few suggestions to ensure legal issues are respected, and hence the site could be trusted. First, one has to familiarize himself with copyright law and its fundamental principles through documents such as Brad Templeton's Guide to common Copyright Myths. Second, the reader has to be vigilant in detecting suspicious websites with questionable content, and actually verifying this content by copy pasting it in search engines such as Google to see if it originates from another source. And finally, users should "avoid fan sites, lyric sites, paper mills, and any site posting newspaper or magazine articles (the full articles, not quotes or links) without also posting explicit permission statements."

The guide also presents five additional questions to consider when choosing a trustworthy website for a particular topic, when a selection of sites is displayed by a search engine [3]:

- Is the origin of the content (i.e. information source) provided by the author?
- Does the site clearly distinguish between opinions and facts when the former is given?
- Is the content of the site relevant to the targeted audience?
- Does the site's perception of the topic or information an objective one, providing certain balance to the reader?

- Is this site similar or comparable to another site that presents information on the same topic?

In examining the principles and the questions explained above, we can state that they are indeed general guidelines and are not based on any form of factual analysis of the website information. The only principle that could be relevant to the trustworthiness, reliability, integrity or fidelity of the facts, is the credibility principle, where the credibility or reputation of the author can be somewhat of an indication of how trustworthy the facts provided really are. However, the last question in the five additional questions suggested by the guide and listed above is directly related to our work, where factual information is compared between sites covering similar subjects. Our work actually aims at measuring this comparison, and eventually translating this measurement into a certain numerical trust value.

In sum, we are inclined to affirm that Ms. Schneider provides one with generic rules that give the user techniques or indicators on how trustworthy the factual content of the site may be, without however giving a method to calculate a numerical value based on a certain scale that would indicate a mathematically based trust level. Some may argue that trust cannot be seen as a rational calculation, and that its assessment needs to be done in an empirical fashion rather than a heuristic one. And this is the very reason why our study will aim at defying this argument, by giving a calculated value to trust which is actually based on our heuristic model (who is in turn based on some empirical data such as a user survey to determine the trusted website), and then comparing these values to the results of our own empirical study (i.e. user feedback method) to justify the effectiveness and accuracy of our method.

As a matter of fact, most litterature relevant to "online trust" does not deal with the factual aspect of the information, as stated previously, but address security and privacy issues. For instance, the works in [5], [6], [7], [8], and [9] identify trust issues by consumers when it comes to engaging in E-commerce. This essentially involves submission of financial and personal information to merchants through the internet. Neilsen [10] gives techniques in order to efficiently communicate trust in web design, once again related to E-commerce web applications. Wang and Emurian [5] take it a step further by giving an overview of several concepts of trust from

different disciplines, which we will explore in the next few paragraphs. They also present "trust-inducing framework interface design features" [5] that could be applied to E-commerce websites in order to optimize trust levels from the customers' perspective. In fact, [5] could be classified as one of the most comprehensive articles covering online trust issues in general, and particularly with regards to e-commerce systems as it reviews most research done so far on the topic. It is for this reason that we will explore this work in further detail.

In order to have a fundamental comprehension of trust, Wang and Emurian summarize several works relating to this concept, and originating from literature in several disciplines including philosophy, psychology, management, and marketing. This allows a clearer definition and conceptualization of the term in general. Here is a summary of how these different disciplines perceive the concept of trust based, for the most part, on the work from Wang and Emurian in [5]:

- Philosophy: Ancient philosophers studied trust in order to depict human nature. They believed that one person trusted another, if he was confident that this person would fear punishment enough to deter him from mischief. In turn, modern philosophy was introduced with a new definition that focused on interpersonal trust. It presented a " ...three-place predicate (A trusts B with valued thing C)..." [5], where "...trustor A needed good judgment to know how much discretion to give. Because trustee B would have had discretionary powers over the entrusted valued thing C, the trustor had to take the risk if the trustee abused the granted power." [5]. Then came political philosophers, who portrayed trust in a social manner, emphasizing values and benefits.

- Psychology: Similar to the discipline of philosophy, most of the psychology literature also concentrated on interpersonal trust. However, this was in parallel to the philosophical definition, as psychologists viewed trust as being specific to one's personal traits; they studied the differences in trust among people according to their individual characteristics, and the effect of these differences with regards to life in general and human interactions. For instance, some psychoanalysts emphasized the impact of "basic trust" [5] on personal characteristics, and how this trust is vital to a "healthy personality" [5]. Also, Deutsch [21] specified the essential role of trust in terms of building a

"cooperative relationship" among humans. Finally, in addition to the definition of interpersonal trust, Rotter in [22] stated that trust is " ...an expectancy held by individuals or groups that the word, promise, verbal, or written statement of another can be relied on" [22].

- Management: Unlike the disciplines of philosophy and psychology who accentuated on interpersonal trust from different perspectives, the discipline of management concentrated on organizational trust, and its importance in actual organizations [5]. Driscoll in [23] defined organizational trust as "...the belief that the decision makers will produce outcomes favorable to person's interests without any influence by the person...". Another researcher identified trust as a "control mechanism" [5] leading to an increased productivity and efficiency among employees working jointly, which in turn leads to a significant improvement in overall "business performance "[5].

- Marketing: The examination of trust in the discipline of marketing was made in a different environment that involved "...distribution channels (e.g. manufacturer – retailer) and buyer-seller relationships." [5] Some researchers indicated how trust predominated business authority or power in terms of increased profits. They also pointed out the importance of trust in helping manufacturers and vendors to achieve "...their full potential..." and "...facilitate long-term commitments..."[5]. Furthermore, the research world concentrated on "relationship marketing" between the seller and consumer. Here, trust was defined as "...a willingness to rely on an exchange partner in whom one has confidence" [24]. This was shown to be capital in the establishment and preservation of long lasting relationships between the two parties, where skills such as "...expertise, likeability, and similarity to customers..."[5] of the salesperson representing the seller were essential in strengthening the "willingness" on the customer's behalf. This in turn allowed the building of trust and reinforced "...the link between the customer and supplier firm" [5].

Wang and Enumiran also enumerated several essential common characteristics of offline and online trust that make the concept easier to understand and grasp. These characteristics were

deduced from the way trust was perceived from previous studies. Table 1 summarizes these characteristics, by describing them briefly from offline and online point of view. As we can see, both offline and online trust share the similar characteristics with some important differences. After examining the table, we can assert that in terms of offline trust, the characteristics seem to be generic, and somewhat reflect the numerous perspectives and definitions of trust, put forth by the disciplines presented in [5] which we have summarized above. Hence, we can deduce that the 'trust level' of our method is better characterized and represented by the offline characteristics, considering that the online characteristics in Table 1 are specific to E-commerce sites. These former traits, in addition to the definitions from the different disciplines (in particular philosophy and psychology), allow us to understand the concept of trust in general, and control certain variables in our research in order to fully develop our method (i.e. with the user in a user survey). However, when analyzing the same characteristics for online trust, we can affirm that they are in relation to trustworthiness of websites that involve certain financial transactions, namely E-commerce websites. In other words, this ensures us that most research done so far in the context of "online" trust, involves e-commerce websites (as stated by Wang and Emurian) , where trust elements and issues are identified, and techniques to handle these issues are suggested in order to enhance online trust from the customer's and the merchant's point of view. Therefore, not only do these academic works come short in providing any significant mathematical model to compute this E-commerce online trust, but they certainly do not address any solutions in order to solve our problem based on trust related to the factual content of a particular website.

Table 1. Summary of characteristics of offline and online trust given by Wang in Enumirian in [5].

Characteristics	Offline Trust	Online Trust
Trustor and trustee	- Two parties in trusting relationship: trusting (trustor) and trusted (trustee). - Parties can be people, organizations, products, etc. - Trust is built if trustee acts in best interest of trustor and on degree of trust trustor puts in trustee.	- Two parties in trusting relationship: trusting (trustor) and trusted (trustee). - Parties are specific entities: e-commerce consumer (trustee) and e-commerce website or merchant (trustor) or Internet itself.

Vulnerability	- Engages vulnerability. - Only an issue in risky or unsure environment. - Trustor has to make himself vulnerable to trustee by risking something important.	- Internet introduces higher risk. - Consumer unsure of degree of risk and consequences during e-commerce transactions. -Consumer not aware of collected involving their activities. -Consumer vulnerable to loss of money and privacy.
Produced actions	- Leads to risky actions. - Action depends on scenario, and involves concrete or intangible assets - Ex.: trust a friend to pay you back; trust your mate to be loyal in relationship	-Leads to specific actions from consumer: - making purchase => providing credit card and personal information - "window shopping" on site
Subjective matter	-Influenced by differences among people, and different contextual factors.	-Also a subjective matter. -Appropriate level of trust for online transactions varies among consumers.

2.2 Reputation

The term reputation is also related to E-commerce or any online business, where the buyer favors a transaction with the seller known to have a good reputation. This concept is clearly illustrated in E-commerce sites such as Amazon.com or eBay.com. Furthermore, reputation has been used in the context of web services. For example, when the composition of web services is needed for an online business operation, the reputation of the individual services or service providers it needs to use is verified. This ensures a highly reputable composition of the service [11]. Also, a web service's reputation is needed to evaluate elements such as privacy issues in web services [12].

There are several models that were introduced to compute the online reputation of business parties and web services. Zacharia et al. [13] presented a collaborative model for computing the reputation of the users involved in E-commerce transactions. The model uses the ratings provided by one user to the other as the primary criteria to compute the reputation. Sreenath and Singh [14] presented an agent-based collaborative approach for computing the reputation of web services. This approach suggests the idea that agents (e.g. service users) will have to work

together in order to independently evaluate service providers. The agents will have to decide on how much weight should be given to each other's recommendations. This method is useful when the collaborating agents have similar needs in terms of service attributes.

A computational model was proposed by Mui et al. [15]. This model focuses on the relationship among trust, reputation and reciprocity. It proposes a probabilistic mechanism for inference among trust, reputation and level of reciprocity in a multi-agent environment (e.g. E-commerce systems). The underlying principle or concept behind this model is that a higher reputation leads to higher trust; where the trust reciprocates actions between two agents; and the reciprocate actions will lead to a higher reputation.

All these methods related to the computation of reputation involve constant feedback of users, whereas our method overcomes this user-feedback dependency.

2.3 Confidence

The term confidence, which is related by definition to the concept of trust (i.e. confidence is defined as a "firm trust" [18]) , has been used in a few different technological contexts. For example, in works [25], [26], [27], and [28] confidence has been used in the data management of sensor network. As for the context of multimedia applications, the term is used in works [29], [30], [31], and [32]. We will give a brief overview of these works and address how they make use of confidence values in sensors/media streams. This will be done by first examining the way we gather the necessary information to determine this confidence and how this information is used to incorporate with streams. Then, we will briefly cover the techniques of merging or fusing these confidence values to determine overall confidence of a group of streams, and whether or not these values were statically or dynamically computed.

In establishing their confidence level, works [26], [27], [30], [31] all use current observation in sensors. However, works [28], [29], and [33] base themselves on the past accuracy in order to determine the confidence level of a stream.

In work [29], a Bayesian formulation adopting logarithmic opinion pool (LOGP) is used as an agreement in assimilating media streams according to their confidence level, with the use pre-established or pre-computed confidence (i.e static confidence). In the work done in [28], Tatbul et al. present the computation of confidence in a stream based on its sensor accuracy in the past. However, they did not use the values of confidence in sensor data fusion, nor did they consider confidence fusion. On the other hand, although the authors of [33] use the Dempster-Shafer theory where of evidence to fuse the confidences, they still employ a pre-confidence or static confidence approach, in contrast with Atrey et al. in [29] who use a more effective dynamic strategy in the computation of confidence in media streams.

Furthermore, from a multimedia perspective, Tavakoli et al. [27] advocate an event detection strategy using historical and spatial information in groups or clusters (current observation in clusters), in establishing a high confidence level in detection. In other words, the amount of sensors and the degree or time that these sensors provide consistent data, are the two main parameters used to determine the confidence level of this sensed data.

In another work, a confidence-based fusion strategy has been used by Iannou et al. [31]. This allowed the combination of multiple feature cues for facial expression recognition, where current features were used in the computation of confidence in a particular cue. The computation of the confidence level in streams from these methods has been based on the past accuracy of the stream, and is computed in a static fashion. These works do not elaborate on the use of the confidence values, and how these values are used in assimilation of information or data. Also, the works lack in making a clear distinction between the spatial and historical information in clusters (i.e. current observation), and the confidence related to it.

Also from a multimedia application context, Atrey et al. [29] use a dynamic approach for the computation of confidence levels in streams based on their agreement/disagreement with trusted streams. In fact, these authors combine the past history of agreement/disagreement and current observation (adopted from other works) in order to compute the confidence level. This agreement/disagreement value between two streams is called the agreement coefficient. This coefficient is determined based on the degree of concurrence or contradiction in the evidence that

streams provide. Also, it is important to note that this agreement coefficient established between two media streams is different that the concept of mutual information between them presented in [34]. The agreement/disagreement coefficient in [29] refers to "...the measure of how agreeing or disagreeing evidences the system obtains based on the two streams" [29], while mutual information in [34] connotes the amount of information conveyed by one stream in regarding another one.

In sum, we can affirm that the work presented by Atrey et al. [29] surpasses the others as it allows a dynamic evolution of the confidence level of streams, instead of a static one, as proposed by the other works with the exception of [25]. Our method is similar to the one stated in [29], but is in context of web information, where the trust level of a web site will be dynamically computed over time based on the similarity measure between the web site in question and a trusted web site within the same domain.

Table 2. A summary of past works related to the computation of trust, reputation and confidence.

The work	Basis of computation	The context
Trust		
Schneider [3]	No computation model.	General guidelines or heuristics for trustworthiness of websites.
Emurian and Wang [5]	No computation model.	General trust definitions from different disciplines. Online trust issues related to E-commerce applications.
Ang et al. [6], Belanger et al. [7], Friedman et al.[8], Fox [9], Neilsen[10]	No computation model.	Online trust issues related to E-commerce applications.
This work	Semantic Similarity.	Dynamic computation of

		trust level for a particular website in a particular domain, based on factual content of the website.
Reputation		
Zacharia and Maes [13]	Ratings provided by one user to the other.	Collaborative- feedback model for *reputation* of a web service.
Sreenath and Singh [14]	Weight of recommendations given by agents (e.g. service users).	Collaborative- feedback model for *reputation* of a web service.
Mui et al. [15]	Probability based mechanism.	Computational model for relationship between *trust* and *reputation* in a multi-agent environment .
Confidence		
Atrey et al. [29]	Past Accuracy - Bayesian formulation using logarithmic opinion pool (LOGP).	Dynamic computation of *confidence* in a media stream.
Tatbul et al. [28] Seigel et al. [33]	Past Accuracy.	Static computation of *confidence* in a media stream.
Frolik et al.[26], Takavoli et al. [27], Luo et al.[30], Ioannou et al. [31]	Current observation - confidence based fusion strategy.	Static computation of *confidence* in a media stream.

Chapter 3

Proposed Method

The proposed method determines the trust in a new non-trusted website using its "Similarity Measure" with the trusted website(s) [35]. The similarity measure between the two websites is computed based on a model named Latent Semantic Analysis (LSA). After stating our problem in section 3.1, section 3.2 provides a brief overview of this LSA model and describes the similarity measure computation in our method. Section 3.3 covers the actual trust computation, being fundamentally the main algorithm and contribution of the thesis work. This section consists of two sub-sections describing the main algorithm: section 3.3.1 which describes the trust computation of a website for a specific domain, and section 3.3.2 which describes the overall trust computation of the website. Section 3.4 examines the architecture of the system, by giving an overview of the System Design: section 3.4.1 gives the system component functionality, while section 3.4.2 demonstrates the main system behavior through a message sequence chart.

3.1 Problem Formulation

Below, we formulate the problem of determining the trust level of a website:

- Let $\mathbf{W} = \{W_1, W_2, \ldots, W_n\}$ represent a set of n websites that we are exploring. Let each website provide information of m domains.

- For $1 \leq i \leq n$, let $0 < \phi^k_{ij}(t) < 1$ be the *similarity measure* between two websites W_i and W_j for a domain k at the time instant t. The term $\phi^k_{ij}(t)$ is determined by using a Latent Semantic Analysis (LSA) [4] document similarity model.
- For $1 \leq i \leq n$, $1 \leq k \leq m$, let $0 < T_i^k(t) < 1$ be the user's trust in the ith website for the kth domain at time instant t. We assume that, for each domain, we have at least one website that we call "trusted". A trusted website is the one that has a trust level greater than a threshold T_{spec}. This trust can be established through traditional user feedback methods, such as a user survey.

Our objective here is to determine the overall trust level $T_i(t)$ of a new "not-so-trusted" website W_i at time t given that its trust level $T_i(t\text{-}1)$ at time $t\text{-}1$ is known. Note that, in absence of any prior information, we assume that the initial trust level of this "not-so-trusted" website is a number close to a positive infinitesimal.

3.2 Similarity Measure Computation

As stated previously, we use the LSA model to measure the similarity between the websites. This method is used to model the semantic similarity between text documents. It has been shown that this model produces inter-rater correlations of about 0.6, which is consistent with human performance in judging semantic similarity between text documents [4]. This was shown to surpass other models such as word-based gram and n-gram, which achieved inter-correlations with human performance of 0.5 each.

LSA starts by forming a $k \times l$ matrix $\mathbf{C} = [c_{ij}]$ where k is the number of words, l is the number of documents in the corpus, and c_{ij} is the frequency of the ith word in the jth document. According to this document representation schema, LSA uses three local weighting functions to measure similarity. The functions reflect essentially the importance of a word within a document, the frequency of a word throughout the whole corpus of documents, and "the number of dimensions retained during the singular value decomposition, which makes assumptions about the complexity of the underlying semantic regularities expressed by the corpus" [4].

These functions are then used to generate a weighted corpus representation **V**= $[v_{ij}]$. This schema is then subjected to singular value decomposition. Afterwards, a variant of the Cosine Model is used to measure the similarity.

For our trust computation, we compute the Semantic Similarity Coefficient (or Similarity Measure) $\phi^k_{ij}(t)$ between two websites W_i *and* W_j for domain k at the time instant t as follows: $\phi^k_{ij}(t)$ is a function that depends on the current similarity value between the websites at time t, denoted as $LSA_{ij}{}^k(t)$, and the similarity measure at time t-1, denoted as $\phi^k_{ij}(t\text{-}1)$. This linear combination model used to find $\phi^k_{ij}(t)$ can be precisely described as follows:

$$\phi^k_{ij}(t) = \alpha \cdot LSA^k_{ij}(t) + (1-\alpha)\cdot \phi^k_{ij}(t-1) \qquad (1)$$

Here, we have $\alpha \in [0,1]$. This is an experimental value used to assign particular weights to the past and current similarity measures, respectively. It is important to note that in the absence of any prior information in equation (1), we assume, for $1 \leq i, j \leq n$, that $\phi^k_{ij}(t)\ (0) = €$ (€= 0.01 in this work). This implies that time $t = 0$ for a particular website, $\phi^k_{ij}(-1)$ is equal to some positive infinitesimal value.

3.3 Trust Computation

3.3.1 Domain-level trust computation

We define $T^k_j(t)$ as the function representing the trust level of a particular "not-so-trusted"website W_j, for domain k at time t. In principle, $T^k_j(t)$ is computed as a function g of two variables:

- $T^k_j(t\text{-}1) \in [0,1]$: the trust level of website W_j for a particular domain k at time t-1;
- $\phi_{ij}{}^k(t) \in [0,1]$: the similarity measure between the trusted website W_i, and the "not-so-trusted" website W_j, at time t.

Precisely,

$$T^k_j(t) = g(T^k_j(t-1), \phi^k_{ij}(t)) \qquad (2)$$

To develop this function, we use an exponential model which allows us to rewrite equation (2) in a more direct form as follows:

$$T_j^k(t) = \frac{T_j^k(t-1) \times e^{\beta(t)}}{N(t-1)} \qquad (3)$$

In the above equation, we make use of two new terms $\beta(t)$ and $N(t)$ which will be described further down. The term $\beta(t)$ represents the growth factor for the trust level T^k_i of website W_i for domain k.

This term $\beta(t)$ is described as follows:

$$\beta(t) = T_i^k(t) \cdot \Delta\phi_{ij}^k(t) \cdot \mu \qquad (4)$$

In equation (4), we notice the appearance of three new terms, $T_i^k(t)$, $\Delta\phi^k_{ij}(t)$ and μ. The first term is essentially the trust level of the so called "trusted" website for domain k, at time t. The second term can be defined as the difference or change in the similarity measures between website W_i and website W_j for domain k at time t. This implies that a higher change in the value of the similarity measure would lead to a higher growth of trust, and vice versa. More precisely, this function can be defined as follows:

$$\Delta\phi_{ij}^k(t) = \phi_{ij}^k(t) - \phi_{ij}^k(t-1) \qquad (5)$$

The last term μ, referred to as the growth value, is also used to control the rate of growth or decay, depending on the fact that the change in the similarity measure, $\Delta\phi^k_{ij}(t)$, is either positive or negative. Hence, μ can hold two distinct values as shown below:

$$\mu = \begin{cases} r_{growth} & \Delta\phi_{ij}^k > 0 \\ r_{decay} & \Delta\phi_{ij}^k \le 0 \end{cases} \qquad (6)$$

The terms r_{growth} and r_{decay} denote the rate of growth and decay in the trust level. Note that the term r_{growth} is also referred to as the growth value.

As for the $N(t)$ used in equation (3), this function is used as a normalization factor to limit the trust level value between [0,1]. Thus, $N(t)$ can be expressed as follows:

$$N(t-1) = T_j^k(t-1) \cdot e^{\beta(t)} + (1 - T_j^k(t-1)) \cdot e^{-\beta(t)} \qquad (7)$$

3.3.2 Overall trust computation

The overall trust $T_i(t)$ for a particular website W_i at a time instant t, can be simply calculated by averaging the trust levels computed for all domains. This can simply be described as:

$$T_i(t) = \frac{1}{m}\sum_{k=1}^{m} T_i^k(t) \qquad (8)$$

Figure 1 below illustrates the high level pseudo code summarizing the proposed method algorithm suggested above.

```
For each Website W_i and W_j;  i and j = 1 to n
        For each Domain k;  k = 1 to m
            Compute Similarity Measure  φ̂_ij(t) between W_i and W_j
            Compute Difference in Similarity Measures  Δφ̂_ij(t)
            Compute Growth Value μ
            {
                IF Δφ̂_ij(t)  > 0
                THEN μ = r_growth
                IF Δφ̂_ij(t)  = 0
                THEN μ = r_decay
            }
            Compute Growth Factor β(t)
            Compute Normalization Factor N(t)
            Compute  Domain Trust Level T^k_j(t),  for website W_j with domain k
        End {2nd For loop}
        Compute Overall Trust Level T_j(t)  for each website W_j
            {
                For each Domain Trust Level T^k_j(t);  k = 1 to m
                    T_j(t)  = SUM(T^k_j(t))
                End {3rd For loop}
            }
End {1st For loop}
```

Figure 1: Algorithm for trust level computation of websites

3.4 System Design

The system is based on a multi-tier, service oriented architecture. Figure 2 gives a general high level overview of the system architecture. Figure 3 shows how the system and its components would be deployed through a conventional deployment diagram, while Figure 4 portrays the interaction or associations between the main or functional components of the system.

3.4.1 Component functionality

- **Client:**
    - The client is a typical web browser (e.g. Internet Explorer, Mozilla Firefox,...). This implies that client requests are simple HTTP requests to a certain web server.
- **Trust Observer**:
    - This component could be seen as a "process controller" responsible for handling the main interactions between the system's functional components. This component controls the communication between the client and the web services, and the communication between the web services themselves, where this process running at the trust observer server is mainly responsible for coordinating the computation of trust of a given web site. This is accomplished by initiating the execution of different web services that use the web sites and trust database.
- **Trust Database**
    - The database stores information such as the mapping of domains to websites, similarity measures, and trust levels.

The Domain Classifier, Parser Process, Similarity Computation, and Trust Computation, are web services responsible for the business logic, and thus justifiably constitute the Business tier. They are the main components of the system. All these essential components have access to the database. Here is a brief description of all the web services' functionality:

- **Domain Classifier Service**
  - Web service that is responsible for mapping or classifying each http requests (e.g. website links) to a particular domain of interest (e.g. health, information technology, international politics, business, entertainment ...).
- **Text Parser Service**
  - Web service that is responsible for extracting the information in the trusted website and/or the not-so-trusted website according to the user's request.
- **Similarity Computation Service**
  - Web service which receives the texts from the Parser (i.e. through the Process Controller and Trust Computation process), and computes the current semantic similarity between two text documents, using LSA.
  - After computing the current similarity measure, the Similarity Process retrieves the previous similarity measure from the database, combines it with the current similarity measure, and computes a new similarity measure.
- **Trust Computation Service**
  - This web service is the entity that executes the trust level computation algorithm, which essentially represents the main contribution of this thesis.
  - The Trust Computation web service receives the similarity measure from the Similarity Process, and computes new trust level based on previous trust level, which is also retrieved from the database.

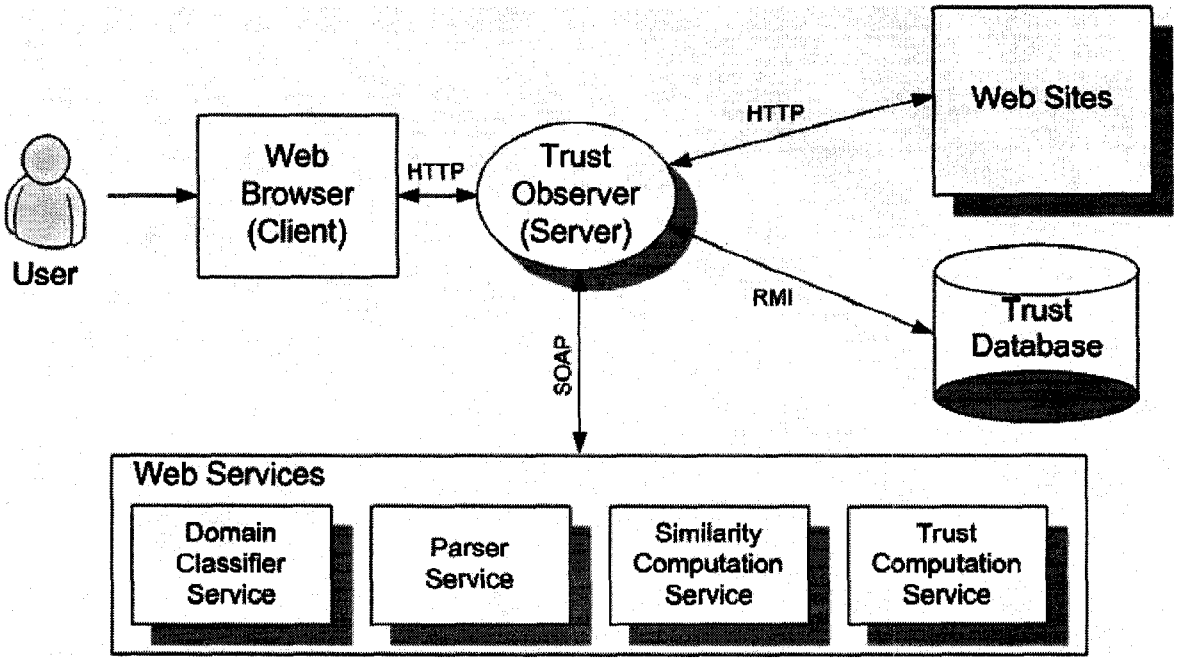

Figure 2: Overall system's architecture

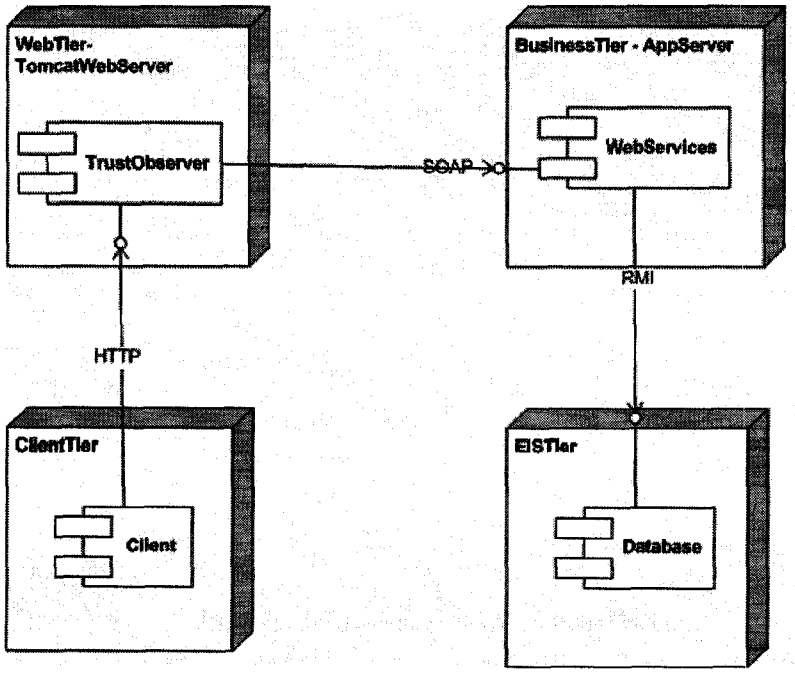

Figure 3: Deployment diagram

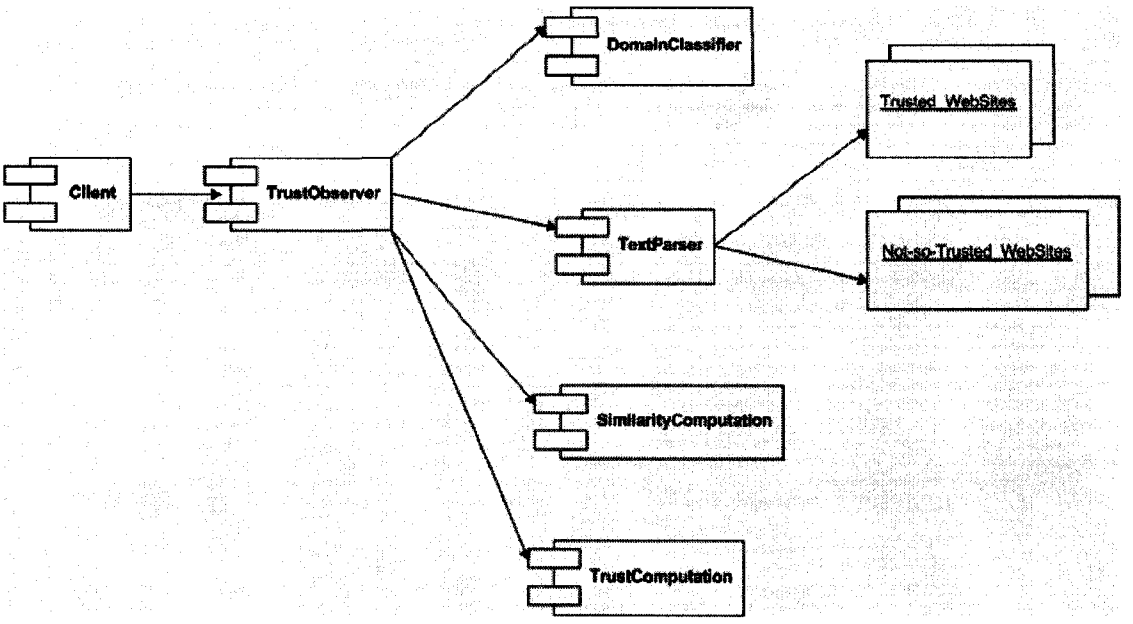

Figure 4: Component diagram

3.4.2 System Behavior

Here are the steps summarizing the main system behavior, or the principal interaction between the system components:

1. Client requests a certain not-so-trusted website through a typical web browser.
2. Domain Classifier process
   2a. Domain Classifier maps the not-so-trusted website to a certain domain.
   2b. Domain Classifier retrieves link of trusted website in the same domain.
3. Parser Process retrieves information from the trusted website and not-so-trusted website.
4. Similarity Computation process
   4a. Similarity Computation web service retrieves previous similarity measure between websites.
   4b. Similarity Computation web service computes current similarity measure between websites.
   4c. Similarity Computation web service computes new similarity measure between websites.
5. Trust Level Computation process
   5a. Trust Computation web service retrieves previous trust level of not-so-trusted website.
   5b. Trust Computation web service computes new trust level of not-so-trusted website.

It is important to note that these steps are all performed in an automatic and dynamic fashion, where the user is completely oblivious to the actual 'behind the scene' process. The user only views the trust level of the website as a numerical value on the actual web page, when this page is actually provided by the browser. Figure 5 illustrates a message sequence chart (MSC) depicting the interaction between components in greater detail.

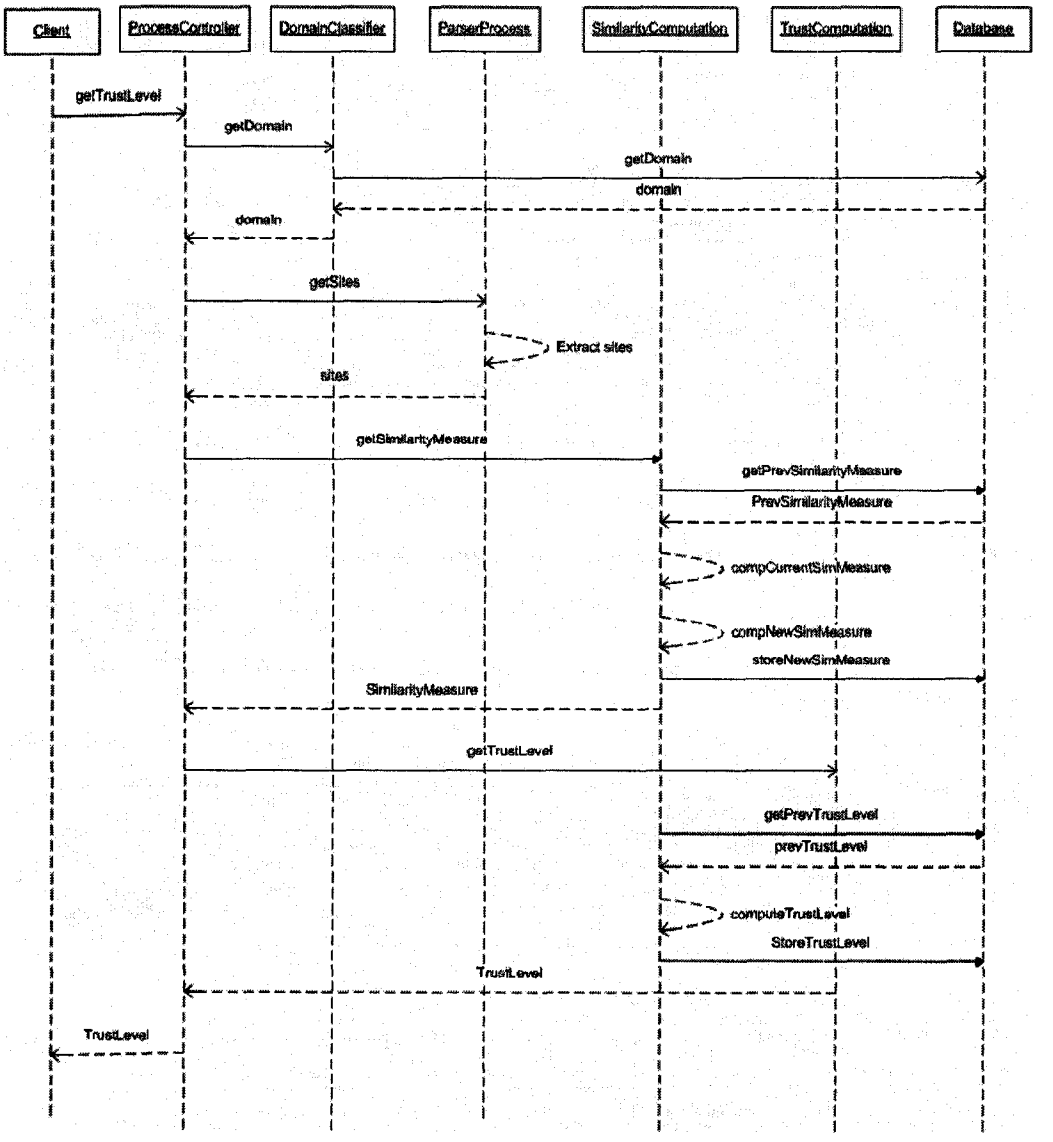

Figure 5: Message sequence chart representing the system's behavior

Chapter 4

Experimental Results

In this section, we present the experimental results in order to demonstrate that the proposed method for dynamically determining the overall trust level of a particular website works well. The experiments have been designed with the following two objectives - first, to demonstrate that the proposed method works well; and second, to establish that it performs as well as any other traditional approach of computing trust with an additional advantage of overcoming the dependency on the user feedback. The organization of this section is as follows. In section 4.1, we describe the data set used in the experiments. Section 4.2 provides an overview of the user survey method used in order to determine the "trusted" website for each domain. Section 4.3 describes the user feedback method employed for the purpose of further evaluating and hence solidifying the results obtained by our method. The next section, section 4.4, presents discussions and analysis of the experimental results. In section 4.4.1, preliminary analysis is conducted allowing us to illustrate dependencies between certain variables in our mathematical model, and hence determine the ideal growth value, an important experimental variable whose value is subjective. Note that the justification of the choice of this value will be demonstrated further on in section 4.4.3. Section 4.4.2 presents the results based on our method, and makes an analytical comparison between these results and the ones attained by the user feedback method in order to justify the strength of our method. Moreover, this section allows us to meticulously examine the results of our empirical user study, in order to comprehend what factors could have influenced users to behave in such a manner, and hence get a clearer understanding of the notion of 'trust level' or 'factual trust' (i.e. on the web) relative to our work. We conclude this chapter

with section 4.4.4, where a consolidated version of the remaining results (from section 4.4.2) are presented. This analysis gives us the opportunity to further validate the utility of our method.

Table 3. Trust levels determined by the user survey methods for each domain.

W_i	Website/ Domain	International Politics	Sports	Health Sciences
W_1	CBC.ca	0.79	0.73	0.63
W_2	CNN.com	0.72	0.68	0.66
W_3	FoxNews.com	0.53	0.56	0.54
W_4	Aljazeera.net	0.67	-	-
W_5	BBC.com	0.82	0.76	0.67

4.1 Data Set

We have used five websites for the evaluation. The websites are the following: CBC.ca, CNN.com, FoxNews.com BBC.com, and Aljazeera.net. These websites have been evaluated for a total of three particular domains: International Politics, Sports, and Health. However, it is important to note that Aljazeera.net is an exception to this, for its evaluation was only done for the first domain due to its lack of information about the last two domains. The data set used in our experiments consists roughly of 90 web news articles for each domain. These articles were gathered over a period of 3 months. The length of the articles varied from 50 to 250 words.

Note that this selection in websites was done in order to cover different media perspectives, from different regions of the globe: CBC, CNN and Fox representing North America; BBC representing Europe; Al-Jazeera representing the Middle-East.

The evolution of trust has been performed based on data collected over a period of three months. Note that some domains such as International Politics had more articles to compare due to the fact that several topics such as the war in Iraq are covered extensively by all websites reporting the news.

4.2 Determining the Trusted Website(s)

A user survey was made in order to establish the so called "trusted" website. Differently put, this survey allowed us to determine the trust levels of every domain for each website, making this website a "trusted" one for that domain. This value is indeed an integral one in our proposed method.

The survey was conducted with 100 users (volunteers) from different disciplines, all being adults. The subjects were asked to rank each website from zero to five (zero being the *lowest* trust level- "distrust" and five being the *highest* trust level- "complete trust") depending on how much the user trusts each website for a particular domain. A typical scenario illustrating this would be the fact that one may trust CBC.com com for International Political news with a rank of three out of five. We averaged the rank values in order to compute the trust level of the "trusted" website for each of these three domains, based on the ratings of the 100 users. The results of this user survey are summarized in Table 3. From these results, we conclude that BBC.com was found to be the most "trusted" website for all three domains. More specifically, this website was found to be trusted 82% when it comes to International Political news, 76% for Sports related news, and 67 % in Health Sciences news.

Note that, a user survey is adequate in such a case where news websites are evaluated, since much of these websites are familiar to the users and their reputation is well known in general. Furthermore, this survey is indispensable in capturing the aforementioned trust-related elements enumerated in section 1.2. The capture of these elements is important in order to integrate the predisposed psychological and social aspects into our method, which ultimately augments its practicality.

However, when it comes to determine the trusted website(s) within a set of specialized websites that concentrate on one particular domain (e.g. medicine, sports, information technology, etc.), typical users may not have an appropriate understanding of the domain in order to make an

accurate trust level or general assessment based on the factual information presented in the website. Therefore, a typical user survey could be unsuitable, and could be replaced by an expert user survey where experts in the domain are consulted based on their knowledge and past experience in that particular domain. We can also consult a renowned trusted source in a particular domain such as the Medical Library Association website (MLA) [26], as mentioned previously, in order to establish the "trusted" websites in a particular domain. Nonetheless, this type of reliable source could be difficult to find for other and somewhat less crucial domains than health or medicine. Also, as mentioned previously, one has to be vigilant in his findings of such sources. In other words, one has to ask himself whether or not this source is actually credible itself, before trusting the information sources it endorses.

4.3 An Empirical User Study

Our second study is entitled "An Empirical User Study" or the "User Feedback Method". In contrast with the user survey, which is primarily needed for the establishment of the "trusted" website followed by preliminary evaluations of our method (e.g. determining dependencies among variables), this method is essential for a more in-depth and further evaluation of our proposed method, in order to reinforce our obtained results. In this way, the comparison between our results and those obtained by the user feedback method will allow us to assess the validity of our proposed method. Evidently, this assessment could not be made without the results obtained by the user survey method, particularly without the "trusted" website. Hence, we can state that both these methods (user survey and user feedback) have an equally important, yet distinctive, complementary role.

The user feedback method was essentially designed in order to view the evolution of trust, and all its related psychological and social factors, with regards to factual information on the internet from a real world perspective. The evaluation consisted of ten subjects, five females and five males, all being graduate students chosen from different ethnicities and disciplines including administration, engineering, science, and social science. Essentially, this type of assortment was chosen purposely in order to cover a wider range of thinking and perspectives, from both genders. Furthermore, the involvement of scholars in this study implicated more knowledgeable

and hence less susceptible individuals, who hopefully will be more vigilant in their readings than the typical and maybe less educated user. The former could be further inclined towards naivety, and consequently find himself to be more gullible and trusting in a shorter period of time. Thus, this variety in cultural backgrounds and school of thoughts would lead to a more uniform overall assessment, and hence more accurate results that depict a better picture of the diversity of our world.

Essentially, the users were given the same articles used in our evaluation, as mentioned above in Section 4.1 explaining the data set. They had to compare each article per domain with the trusted article: first, the users had to judge the similarity of these documents on a ten-point scale (from zero to ten; zero being "completely unrelated" and ten being "completely related"); second, they had to give a trust level value also on a ten-point scale (zero being the *lowest* trust level- "distrust" and ten being the *highest* trust level- "complete trust").

Moreover, users were given each article from the "not-so-trusted" sources, without giving the name of the website. In other words, each daily article was identified by a number which referred to the same anonymous website throughout the entire experiment (e.g. all articles identified with a "2" came from CNN – but this website remained anonymous to the users). Furthermore, the subjects were asked to give trust levels based not only on the current similarity measure between the "trusted" and "not-so-trusted" article, but based on previous similarity measure given by the user, and most importantly, based on the previous trust level given by the user to that particular anonymous source. These previous values were provided by us to the user. Note that the name of the established trusted website (i.e. from the user survey) was also kept anonymous from the users who participated in the empirical study.

All this allows an evolution of trust from the user's perspective, without biasness towards a particular site since the user only identifies these "not-so-trusted" websites by numbers without knowing their names. In this way, a user could start to trust a particular site, without knowing the name of the site or any prior information about that site, which could influence his decision. Thus, we hope to ultimately reach different conclusions that will help not only in justifying the effectiveness of our method if used in the real world, but will also help us in seeing whether or

not we have clearly understood, defined, and analyzed this notion of internet "trust" based on factual information as we have discussed throughout this thesis. Simply put, this empirical study will allows us to assess the strength of our method by further understanding the psychological dimension of how typical users trust or learn to trust the factual content presented in a particular website.

4.4 Test Cases

We have designed the following test cases as a part of the experiments:

Experiment 1. This experiment is designed to show the dependency and correlation between the trust level of a not-so-trusted website with respect to its similarity measure with the trusted website over a period of 90 days. Note that in this experiment a growth value of 3.5 (r_{growth} = 3.5). Since the choosing of such value is subjective, the rationale behind this decision is explained in Exp#3.

Experiment 2. The main purpose of this extensive experiment is to examine and analyze the results (e.g. similarity measures, trust levels,...) obtained from our empirical user study by comparing them to those obtained from our proposed method, in order to verify its' practicality and effectiveness. Moreover, this analysis will allow us to assess certain statements made with regards to the notion of "trust" in general, by looking at psychological and social facets that could explain the results obtained from the empirical user study. Once again, the choice of the growth value is 3.5, a decision which will be justified in Exp#3.

Experiment 3. This experiment would show how the growth of the trust level of the "not-so-trusted" websites is affected by different values of r_{growth}, and hence justify the rationale of using such a particular value over another, as was done in Exp#1 and Exp#2. Differently said, this experiment is essentially designed to determine the ideal growth value, given the circumstances.

Experiment 4. This experiment is an extension of the second experiment. It will depict a consolidated version of the results already covered in Exp#2 in order to further solidify the

strength of our proposed method. Also, additional analysis will be given with respect to the results obtained by the user feedback method, to help us understand our method's limitations and further enhancements.

We describe these four experiments in the subsequent sections.

4.4.1 Experiment 1 – Trust Level vs. Similarity Measure

In this experiment, we study the dependencies among important variables in our proposed mathematical model. In particular, we will examine the relation between the trust level of a particular "not-so- trusted" website, and the similarity measure between this website and the "trusted" website.

Figure 6 depicts a graph which shows the behavior of the evolved trust level in a website W_2 (CNN) for the first domain, in terms of its similarity measure with the trusted website W_5 (BBC) in that same domain. In this graph, the x-axis denotes the timeline in days, whereas the y-axis represents the trust level from 0 to 1. The first curve observed, denoted by $T'_{2,}$ represents the trust level computed by our method of a not-so-trusted website $W_{2,}$ whereas the second, denoted by $\phi'_{2,5,}$ portrays the semantic similarity measure between the former website and the trusted website. Moreover, a growth value r_{growth} of 3.5 was used here in order to compute the trust level. The value selection of this subjective variable will be justified further on, in Section 4.3.3 (Experiment 3).

When examining Figure 6, it clear that an increase in the similarity measure between the two websites, leads in turn to an increase in the trust level. Conversely, a decrease in trust level is caused principally by a decrease in the similarity measure. This is an apparent implication of dependency and correlation between both variables, where the trust level is in effect proportional to the similarity measure.

Let us illustrate the importance of this proportionality by observing both curves in Figure 6 within specific time frames. We observe a sharp increase between days 0 -10 and days 55-59, in the similarity measure, which in turn leads to an increase in trust level within the same period of time. This is a result of a high semantic similarity obtained between the compared news articles (from the "not-so-trusted" and "trusted" websites) over these periods of time. In contrast, a sharp decrease in similarity on days 52-53, causes a corresponding decay in trust level at that time.

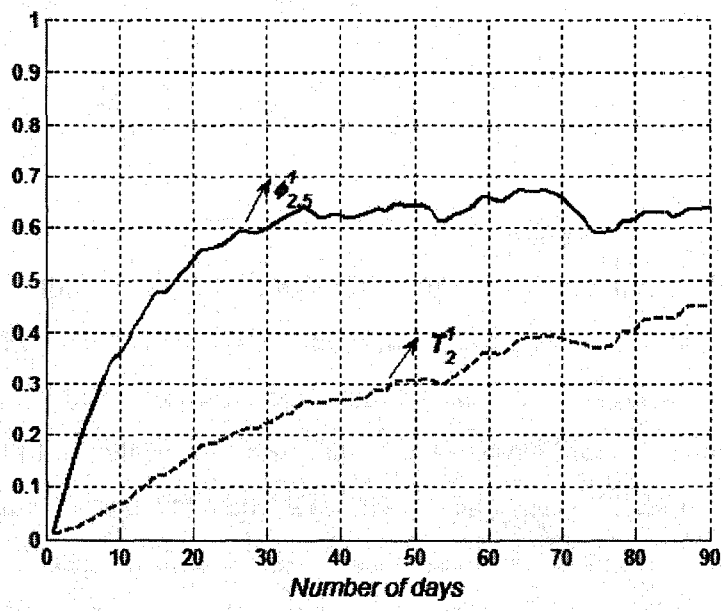

Figure 6. Trust level versus similarity measure over a period of time.

Furthermore, we can see from Figure 6 that both curves in the graph increase/decrease different rates. More precisely, both similarity measures and trust levels increase at a higher rate during the early stages of the study, and then seem to increase at a slower rate. Hence, we can state that higher similarity measures, who had lead to higher increases in trust level in the beginning, lead to slightly lower increases towards the end.

The next section is indispensable to this entire work. It will present the results of our empirical user study, and compare these results to the ones obtained by our proposed method. This will allow us to corroborate the usefulness and effectiveness of our proposed method.

4.4.2 Experiment 2 – Our method versus Empirical study: An example

After assessing from our previous investigation that trust is indeed an evolving process built over a certain period of time, a fact which is in accordance to our proposed method, we can state that the need of this case in our empirical study is indeed most essential to justify the practical utility and effectiveness of our method.

Figure 7 is separated in two parts or two graphs, showing two sets of curves. In Figure 7(a), we have curves $S^1_{1,5}$ and $\phi^1_{1,5}$, representing the similarity measures between the first website and the trusted website, using the user feedback method and our proposed method, respectively. Note that $S^1_{1,5}$ represents the average semantic similarity values by the ten users. In Figure 7(b), we see curves $E_1{}^1$ and $T_1{}^1$, which respectively represent the average trust level values of the users based on the first case of our empirical study, and the trust level values based on our method. Both sets of curves are plotted using data collected over a period of 90 days.

Note that these curves represent similarity and trust values for the first news domain (International Politics) of the first website (W1- CBC.ca). Furthermore, the two remaining news domains (Sports and Health) where not represented graphically in order to prevent redundancy, for the result trends were very similar to those of the first domain. This redundancy of trend also explains the reason why the detailed analysis for the remaining 'not-so-trusted' websites (W2-W4) was also omitted. In order to confirm this highly related tendency among domains and websites, the similarity in the results will be presented further in this section, using a briefer format.

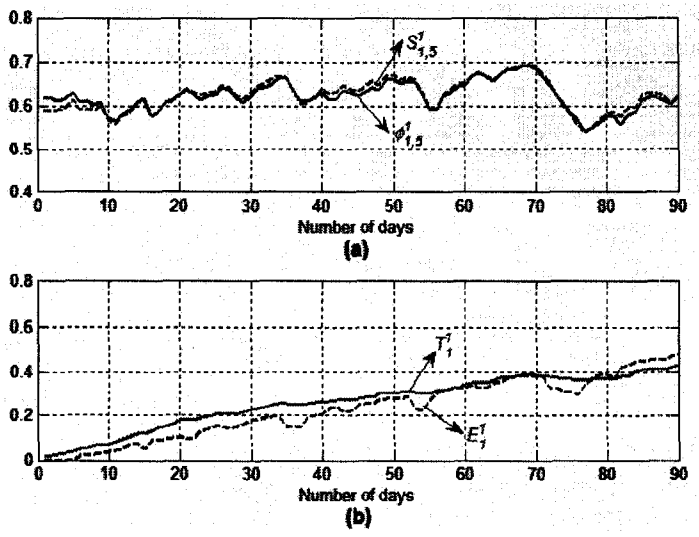

Figure 7. Trust levels and similarity measures according to the empirical study and to our proposed method.

We can see from the graph shown in Figure 7(b) that both curves behave in a similar manner, where a progressive evolution of the trust levels is observed. In addition, the trust level values increase/decrease at a similar rate over the period of time, and both curves have some intersection points (e.g. day 58, 70, 81, etc.), where trust level values between both methods are identical or very similar. For instance, between days 6-10 there is a sharp increase in trust levels, and between days 29-34 a similar increase is observed due to a higher observed similarity between W1 and W5 during the same period of time. This higher similarity can be seen in the graph above (Figure 7(a)) at the corresponding period of time. In contrast, between days 35-37 or days 53-54 for example, a decrease is perceived, which could be explained by similar yet opposite reasons as stated previously: there was low similarity values observed by the users and computed by our LSA model at that point in time, which consequently affected the value of trust. Moreover, by examining the graph in Figure 7(b), we see that user's trust level is slightly lower than the trust computed by our method for most of the time, until day 57 approximately, where the trust value based on the empirical study was very similar to the trust level dynamically computed by our method. Afterwards, around the 75th day, the users' trust surpassed that of our method. This entire trend could be explained for several reasons, which is discussed in the next paragraphs.

From a social and psychological perspective, we can assume that these educated, and fairly more skeptical or less vulnerable users, were somewhat reluctant to trust these unknown websites. This unwillingness to fully engage remained persistent for a period of approximately two months, even when documents were found to be quite similar to those coming from the trusted website. However, after this period of time, the users found themselves more familiar with the sources. Consequently, their skepticism was somewhat lowered, and their readiness to engage eventually augmented, resulting in an increase in their trust.

For example, if a user gave a trust level of '2' on certain day, all the other trust level values given after that day were infrequently lower than '2', even if the user found lower similarity measures between the trusted and 'not-so-trusted' websites. Instead, users gave this trust level value repeatedly for a longer period of time, leading to a much slower increase in the trust level. This behavior was particularly observed after a several weeks within the empirical study, where lower similarity values did not directly affect the trust level values as much, since users had already built a certain trust level with the unknown sources. For instance, an average decrease in similarity from a value of 7 to 6, did not necessarily lead to a decrease in the average users' trust level, whereas a decrease in the same manner according to our method, actually lead in turn to a decrease in trust level.

On the other hand, in the cases where trust level values were exceedingly decreased by a particular user based on lower similarity measures observed, most users returned, after a short period of time only, to the initial trust level given, as soon as the current similarity measure observed was considered higher than the value that caused the decrease in the trust level, even if this current similarity measure was not necessarily amongst the highest observed by the user. Differently put, users had to observe lower similarities over several consecutive days in order to cause them to significantly lower their previously indicated trust value. However, when the users returned to their initial trust level value given before the sharp decrease, this value was kept for a "longer than expected" period of time, which in turn led to an overall slower increase of users' trust. It is essential to note that this noticeable decrease in trust values at certain times (e.g. on days 35-37, 53-54, 72-76) by users was not directly proportional to the decrease observed by our method since the scales used for each method was not identical, limiting the

users' assessment. For example, when a decrease in trust by a user went from 4 to 3.5 at time *t*, our method saw a decrease from 4 to 3.9 at the same time; a difference of 0.4 between both methods. This type of difference is almost inevitable when there is use of empirical studies involving users, and it would seem to be quite unreasonable to ask users to start using a tenth of a point on a ten-point scale to make either trust or similarity assessments!

The behavior discussed in the above could also be interpreted from another psychological point of view, where most users were not able to assess or attribute an exact numerical value that one should add or deduct to the previous trust level, based on the similarity measure. Otherwise stated, the mental relation between similarity and trust was difficult to establish in terms of the limited ten-point scale used for the evaluation.

From a different viewpoint, differences between trust level values from the two methods could have been affected from the fact that several pairs of news documents were found to have a slightly higher/lower semantic similarity by the users than the similarity measure calculated by the LSA model (i.e. depicted in Figure 7(a)). For instance, several articles covered the main idea for the most part, but drifted off in a different direction for the remainder of the article, while the main message was still conveyed in a similar manner, with some aspects varying. This facet was not captured effectively by the LSA model, which gave such pair-wise comparisons a lower similarity value than the users, and human judgment was needed to make a more adequate detection of semantic similarity. On the other hand, some articles reported similar details, but the message or main theme of the article was not essentially the same. Also, users might have seen certain bias elements that could have influenced the articles' credibility even when giving factual information in coherence with the trusted website. Here's a real-life scenario to illustrate this point: let's say that one reads an entire article conveying several known facts consistent with those from a credible source, and then finds one important and subtle fact, not found or contradicting the well-known source. This one element could indeed jeopardize the entire articles' credibility in the reader's eyes, even if the remainder of the article is actually very similar to the credible source.

In summary, despite the slender discrepancy among growth rate between the two methods and the limitations mentioned above, we can assert, according to Figure 7(b), that the overall trend between the methods is quite similar. Hence, this analysis allows us to deduce that both methods are quite comparable, justifying the practicality and effectiveness of our method when used in the real world. This statement will be further validated in Section 4.4.4, where the mean square difference between the two methods will be presented for all the domains in all the websites.

The next determine will help us justify the use of an important value in the main algorithm of our proposed method. This variable is called the growth value. Its value was particularly used in the current experiment in order to compute the trust level based on our method.

4.4.3 Experiment 3 – Determining the Ideal Growth Value

In this experiment, we aim to determine the appropriate or ideal growth value, r_{growth}, an important element of the growth factor function $\beta(t)$ (Equation (4) in Section 3.3.1) in the main algorithm of our proposed method. Essentially, this subjective value mainly dictates the rate of growth in the trust level. It also indicates the rate of decay or decrease (denoted by r_{decay}), for this value is inversely proportional to the growth value.

The 'appropriate or 'ideal' growth value is one where a not-so-trusted website would be able to achieve a trust level close to what it obtained using the user feedback based method. We define a term D^k_i as the mean square difference between the trust level of a website W_i for domain k when our method and the traditional user feedback method are used. In short, we refer to this term, D^k_i, as the individual difference ('individual' implying one domain within one website) in trust levels between our method and the user feedback method. The term D^k_i is computed as follows:

$$D_i^k = \frac{1}{l}\sqrt{\left(\sum_{t=1}^{l} \left| (T_i^k(t))^2 - (E_i^k(t))^2 \right|\right)} \tag{9}$$

Note that in equation (9), $E_i^k(t)$ is the trust level of the i^{th} website for the k^{th} domain at time t using the user feedback method. Also, the term l represents the total length of time that the summation of squared differences between trust levels (inside the squared root) is done over. In our case, based on the empirical user study, this summation is done over a period of 90 days (i.e. $l = 90$).

Equation (10) gives the average difference, D_i, for all domains of a particular website W_i. It is computed as the summation of all individual differences, D^k_i, for a particular domain, divided by the total number of domains, m:

$$D_i = \frac{1}{m}\sum_{k=1}^{m} D_i^k \tag{10}$$

Finally, the overall average difference D_{all} for all the domains in all the websites, between our method and the feedback based method is computed as the summation of all average differences, D_i, divided by the total number of websites, n. The mathematical equation is as follows:

$$D_{all} = \frac{1}{n}\sum_{i=1}^{n} D_i \tag{11}$$

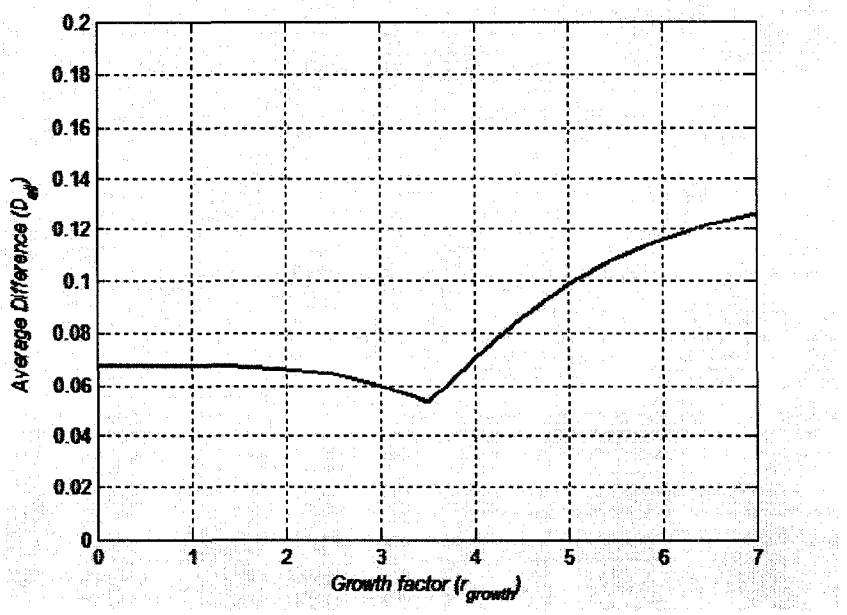

Figure 8. The overall average difference in trust levels, between our method and the user feedback method, when different growth values are used.

The overall average difference, D_{all} (for $1 \leq i \leq n$), is plotted for different growth values ($1 \leq r_{growth} \leq 7$) in Figure 8 above. This is done in order to examine the value of r_{growth} at which the value of D_{all} is minimal. Note that the value of r_{decay} used to compute the trust levels for each domain (and hence compute D^k_i followed by D_i), is 20% of the growth value ($r_{decay} = 0.2\ r_{growth}$). This decay value, being inversely proportional to the growth value, has an opposite yet equally important role as the latter: it dictates the rate of decrease when the difference in similarity measures, $\Delta\phi$, (i.e. $\Delta\phi$ being the difference in similarity measures between two consecutive days, t and t+1), is negative.

As can be seen in Figure 8, the overall average difference D_{all} decreases as the growth value increases. This is observed until a certain point where the lowest difference is reached when the growth value (r_{growth}) has an approximate value of 3.5. At this point in time, the difference is minimal with a value very close to 0. This implies that the trust levels dynamically computed according to our evolving model are 'very similar', in numerical value, to those given by the

users in our empirical user study. Note that this value is in fact the one used in the first and second experiment. This general tendency could be confirmed for each domain in each website, for the average individual differences among these domains and websites is indeed the value represented and reflected by this overall average difference in Figure 8. In other words, the curve illustrated in Figure 8 clearly reflects the overall behavior of the curves representing each individual difference D^k_i versus different growth values (i.e. ten graphs of D^k_i vs. r_{growth}) if these graphs were to be plotted. In turn, this could imply that the ideal value of the growth factor, where the dynamically computed trust level values are closest to the users' trust values, is approximately 3.5 for this particular evaluation period of 90 days. Put in a practical sense, we would have to use a growth factor of 3.5 in our method in order to reach similar trust level values established by the users in the empirical study, given the results obtained (i.e. similarity measures obtained by our LSA model) by our method for this 90 day period.

After this growth value of approximately 3.5 is reached, the value of the difference between the methods increases sharply, as expected. This growth is made at a faster rate than the decrease initially observed until the growth value reaches 3.5. Eventually, after a certain growth value, this difference will follow a more gradual increase.

To conclude this experiment, we can state that the subjective decision making of the growth value could be seen as another overhead to our proposed method. However, we have shown that the ideal growth value could be tailored to make the trust levels provided by our method more realistic and closer to the values established by such techniques as empirical user studies. With a period of three months, the results obtained seemed quite reasonable and practical. This allowed us to adapt our method in such a way as to reflect how factual trust is actually built in users' minds. Nonetheless, a longer empirical study involving more subjects (e.g. 100 users) would without a doubt lead to better results.

4.4.4 Experiment 4 – Overall Result Verification

This experiment was designed to further corroborate the utility and practicality of our proposed method.

We present the consolidated results of the mean square difference, D^k_i, for each domain in each not-so-trusted website in Table 5 below. As we can see, the differences for the same domains are comparable between websites, indicating a similar performance in trust levels among the different websites.

Table 4. Mean square difference, D^k_i, between our proposed method and the user feedback method for each domain (*D1-D3*) in each website (*W1-W4*).

	W_1	W_2	W_3	W_4
D_1	0.04320	0.04522	0.06827	0.06183
D_2	0.02025	0.02013	0.02045	-
D_3	0.07568	0.07541	0.08984	-

For instance, we observe noticeably lower differences in the sports domain (row D_2 in Table 5), for all the websites, than any other domain. This is due to the fact that general sports information is mainly based on well-known facts displayed across most news sites. In fact, most sports articles chosen for this study reflected this general and commonly found pool of sports information, where game reports revealing scores, player injuries, summaries of press conferences, etc. were the main topics covered. However, there is in effect more complex sports information comprised of intricate facts that could be questionable (e.g. player trade rumors, sales of sports teams, sports scandals, etc.) found in sport specific websites. Perhaps the difference between the two methods would have been higher if articles were taken from such sites.

Conversely, a relatively higher difference among the health domains (row D_3 in Table 4), is observed. This tendency could be explained by a certain difficulty in the finding of relevant articles (i.e. articles relevant between each other) for comparisons in the study due to the limited health information provided by general websites, and the breadth of the topic. Therefore, the information found in health articles for this study did not present as high similarity measures as the sports domain or even the political news domain, for topics covered in several articles were somewhat similar, but presented from different angles with a variance in details. More importantly, this variation lead to wider difference in semantic similarity assessments between users and our LSA model, which in turn lead to a more noticeable difference in trust levels. Nonetheless, the use of health or medically oriented websites, experts in the domain, would have allowed us to choose more specific and relevant topics, and perhaps get better similarity results among articles leading to higher trust levels. But then again, in the cases of extreme specificity, certain subjects are covered by a few sources only, or there are several disagreements or contradictions with regards to this particular subject, making it difficult to not only make adequate comparisons due to the limited information, but to find a trusted source in this very specific sub-domain. This could be seen as another, less conspicuous, limitation to consider with our proposed method; a limitation derived from the previously mentioned restraint (in Section 4.2) of finding appropriate and sufficient techniques and/or resources to determine the trusted website in multifaceted domains with various sub-domains.

In summary, according to the results illustrated above in Table 4, we can affirm that the individual differences between the two methods are very small for all domains in all of the websites. Thus, we can conclude that our proposed method is quite comparable to the user feedback method. In fact, this shows that our method can be employed in a real life scenario in order to build the trust level of several websites for numerous domains. Hence, users would be allowed to view web pages with their respective trust level, as our method could be integrated with pre-existing systems such as search engines, where search results would be based not only hit counts, but on trust levels also.

Chapter 5

Conclusions and Future Work

In this thesis work, we have presented a method to dynamically compute and evolve the trust level of a not-so-trusted website for a particular domain (e.g. politics, health, sports, economy, information technology,...) based on how similar its content is with a trusted website of the same domain. This 'trust level' in informative websites allows us to determine the trustworthiness and credibility of the factual information presented in a "not-so-trusted" website, in contrast to the more familiar concept of 'internet/online trust', which involves security and cryptography issues in E-commerce sites.

As shown in the experimental results, the proposed method allows us to surmount the users' feedback dependency - which is typically used as an essential element in the computation or depiction of online trust, reputation and confidence - by dynamically computing the trust level of a website without the need of constant feedback from users to perform this computation. This statement is reinforced after evaluating our method with respect to the user feedback method/empirical user study – a study allowing us to examine from a psychological and social perspective the progressive and evolving aspects of trust in user's minds. As the results obtained from both methods are quite comparable, we can assert that our proposed model to dynamically compute the factual trust of a website is indeed a practical and efficient one. Although the user survey is somewhat essential to determine the so called "trusted" website, the proposed method has an added advantage of allowing us to compute the trust level of the other websites whose trustworthiness is not yet available.

This being said, our proposed method gives not-so-trusted websites the opportunity to increase trust in their users and eventually become comparable to the trusted website, or any trusted websites in a particular domain. These not-so-trusted and often less popular websites are frequently overlooked due to a certain bias element pre-encrypted in the user's mind for one reason or another, or due to a limited budget and inability to join forces with big names on the Web in order to help them become more marketable, commercial, and hence more visible and trustable to the users. Our method aims at reducing the effect of these discriminative elements and this lack of public awareness, and ultimately help not-so-trusted websites become more trusted. From a different perspective, our proposed model can also be viewed as a way to expose more 'popular' websites which present information that is not as trustworthy as it seems.

Although our proposed method's utility is quite evident, its performance is limited by certain integral components such as the Latent Semantic Analysis (LSA) model used to compute the semantic similarity between text documents, and our text parser utilized in the automated editing of texts from websites. Nevertheless, with the ongoing and extensive research in text processing, including natural language understanding and lexical semantics, which could potentially lead to more accurate techniques in modeling semantic similarity between texts and much improved text parsing algorithms or programs, our method's performance would be certainly enhanced and its practicality would become of clearer evidence.

Future work would include a different or less biased method in order to determine the trusted website. For instance, impartial experts in each domain could be consulted in order to obtain their trust level in a particular website based on their personal experiences combined with collected data from this research or other research. From a different, yet related, approach, our proposed method could be fine-tuned to accommodate certain ethnicities, communities, or genders for example. In this case, the trusted website would be determined according to the specific views, standpoints, and general cultural background of to this particular group of people.

Finally, other future work consist in the extension of our method in order to compute the trust level of websites that have other media content such as images, video and audio. This trust level

computation would be performed by finding the semantic similarity between multimodal websites.

Bibliography

[1] Brin, S. and Page, L. 1998. The anatomy of a large-scale hyper textual web search engine. *Computer Networks and ISDN Systems* 30(1–7), 107–117.

[2] Medical Library Association. URL:http://www.mlanet.org/resources/medspeak/topten.html.

[3] Schneider G. K. Beyond Algorithms: A librarian's guide to finding web sites you can trust. URL:http://www.google.com/librariancenter/articles/0601_02.html.

[4] Lee, M. D., Pincombe, B., and Welsh, M. 2005. An empirical evaluation of models of text document similarity. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*. Austin, USA, 1254–1259.

[5] Emurian, H. H. and Wang, D. Y. 2005. An overview of online trust. *Computers in Human Behavior 21*, 105-125.

[6] Ang, L., Dubelaar, C., and Lee, B.-C. 2001. To trust or not to trust? A model of internet trust from the customer's point of view. In *Proceedings of the 14th Bled Electronic Commerce Conference*. Bled, Slovenia, 40-52.

[7] Belanger, F., Hiller, J. S., and Smith, W. J. 2002. Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *Journal of Strategic Information Systems 11*, 245–270.

[8] Friedman, B., Howe, D. and Kahn, P. 2000. Trust online. *Communications of the ACM 43(12)*, 34–40.

[9] Fox, S. 2000. Trust and privacy online: Why Americans want to rewrite the rules. The Pew Internet & American Life Project, Washington, DC. URL: http://www.pewinternet.org/reports/pdfs/PIP_Trust_Privacy_Report.pdf.

[10] Neilsen, J. 1999. Trust or bust: Communicating trustworthiness in web design. Jacob Nielsen's Alertbox. URL: http://www.useit.com/alertbox/990307.html.

[11] Liu, Y., Ngu, A. H. H., and Zeng, L. 2004. QoS computation and policing in dynamic web service selection. In *Proceedings of the ACM Conference on World Wide Web*. New York, USA, 66–73.

[12] Rezgui, A., Bouguettaya, A., and Malik, Z. 2003. A reputation-based approach to preserving privacy in web services. In *Proceedings of the International Workshop on Technologies for e-Services*. Vol. LNCS 2819. Hawaii, USA, 91-103.

[13] G. Zacharia, A. M. and Maes, P. 1999. Collaborative reputation mechanisms in electronic marketplaces. In *International Conference on System Sciences*. Hawaii, USA.

[14] Sreenath, R. M. and Singh, M. P. 2004. Agent-based service selection. *Journal on Web Semantics 1(3)*, 261279.

[15] Mui, L., Mohtashemi, M., and Halberstadt, A. 2002. A computational model of trust and reputation. In *International Conference on System Sciences*. Hawaii, USA, 280–287.

[16] Trust and Trustworthiness on the Web. URL: http://www.audiencedialogue.org/trust.html.

[17] Consumer and Patient Health Information Section. URL: http://caphis.mlanet.org/.

[18] 1971. Oxford English Dictionary: The Compact Edition. New York: Oxford University Press.

[19] Husted, B. 1998. The ethical limits of trust in business relations. *Business Ethics Quarterly 8(2),* 233-248.

[20] Nissenbaum, H. 2001. Securing trust online: Wisdom or oxymoron? *Boston University Law Revie 81*, 101–131.

[21] Deutsch, M. 1962. Cooperation and trust: Some theoretical notes. *Nebraska Symposium on Motivation 10*, 275–318.

[22] Rotter, J. B. 1967. A new scale for the measurement of interpersonal trust. *Journal of Personality 35*, 651–665.

[23] Driscoll, J. W. 1978. Trust and participation in organizational decision making as predictors of satisfaction. *Academy of Management Journal 21(1)*, 44–56.

[24] Dwyer, R. F., Schurr, P. H., and Oh, S. 1987. Developing buyer–seller relationships. *Journal of Marketing 51(3)*, 22–27.

[25] Hughes, K., and Ranganathan, N. 1993. A model for determining sensor confidence. In *IEEE International Conference on Robotics and Automation*. Vol. 2. Tampa, FL, USA, 136-141.

[26] Frolik, J., Abdelrahman, M., and Kandasamy, P. 2001. A confidence-based approach to the self-validation fusion and reconstruction of quasi-redundant sensor data. In *IEEE transactions on Instrument and Measurement* 50(6), 1761-1769.

[27] Takavoli, A., Zhang, J., and Son, S. H. 2005. Group-based event detection in undersea sensor networks. In *Second International Workshop on Networked Sensing Systems*. San Diego, California, USA.

[28] Tatbul, N., Buller, M., Hoyt, R., Mullen, S., and Zdonik, S. 2004. Confidence-based data management for personal area sensor networks. In *The Workshop on Data Management for Sensor Networks*, 24-31.

[29] Atrey, P. K., Kankanhalli, M. S. and El-Saddik, A. E. 2007. Confidence building among correlated streams in multimedia surveillance systems. In *Proceedings of the 13th International Conference on Multimedia Modeling*. Singapore, 155-164.

[30] Luo, J. and Boutell, M. 2004. A probabilistic approach to image orientation detection via confidence-based integration of low-level and semantic cues. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Washington, DC, USA.

[31] Ioannou, S., Wallace, M., Karpouzis, K., Raouzaiou, A., and Kollias, S. 2005. Confidence based fusion of multiple feature cues for facial expression recognition. In *The 14th IEEE International Conference on Fuzzy Systems*. Reno, Nevada, USA, 207–212.

[32] Yu D. and Frincke, D. 2005. Alert confidence fusion in intrusion detection systems with extended Dampster-Shafer theory. In *The 43rd ACM Southeast Conference*. Kennesaw, GA, USA.

[33] Seigel M. and Wu, H. 2004. Confidence fusion. *in IEEE International Workshop on Robot Sensing*. 96-99.

[34] Conaire, C. O., Connor, N. O., Cooke, E., and Smeaton, A. 2006. Detection thresholding using mutual information. In *International Conference on Computer Vision Theory and Applications*. Setubal, Portugal.

[35] Ibrahim, H., Atrey, P.K., and El- Saddik, A. E. 2007. Semantic similarity based trust computation in websites. The 1st ACM International Workshop on The Many Faces of Multimedia Semantics. Augsburg, Germany, 65-72.