Trust on the World Wide Web: A Survey

Jennifer Golbeck

University of Maryland
College Park
MA 20742
USA
jgolbeck@umd.edu
Foundations and Trends® in Web Science
Volume 1 Issue 2, 2006
Editorial Board

Editor-in-Chief:
Wendy Hall
University of Southampton
wh@ecs.soton.ac.uk
Nigel Shadbolt
University of Southampton
nrs@ecs.soton.ac.uk

Editors
Tim Berners-Lee (MIT)
Lorrie Cranor (Carnegie Mellon University)
Dieter Fensel (DERI)
Carole Goble (University of Manchester)
Pat Hayes (IHMC)
James Hendler (University of Maryland)
Arun Iyengar (IBM Research)
Craig Knoblock (USC)
Ora Lassila (Nokia Research)
Cathy Marshall (Microsoft)
Ben Shneiderman (University of Maryland)
Danny Weitzner (MIT)
Yorick Wilks (University of Sheffield)
Editorial Scope

Foundations and Trends® in Web Science will publish survey and tutorial articles in the following topics:

- Agents and the Semantic Web
- Content Management
- Databases on the Web
- Data Mining
- Dependability
- Emergent behaviour
- Human-Computer Interaction
- Hypertext/hypermedia
- Languages on the Web
- Mobile/pervasive
- Multimedia
- Network infrastructures
- Performance
- Scalability
- Security
- Semantic Web — metadata, ontologies
- Standards
- Trust and Provenance (Policy)
- Universal Usability
- User Interfaces
- Web Searching/Information Retrieval
- Web Services

Information for Librarians

Foundations and Trends® in Web Science, 2006, Volume 1, 4 issues. ISSN paper version 1555-077X. ISSN online version 1555-0788. Also available as a combined paper and online subscription.
Trust on the World Wide Web: A Survey

Jennifer Golbeck

University of Maryland, College Park, MA 20742, USA, jgolbeck@umd.edu

Abstract

The success of the Web is based largely on its open, decentralized nature; at the same time, that allows for a wide range of perspectives and intentions. Trust is required to foster successful interactions and to filter the abundance of information. In this review, we present a comprehensive survey of trust on the Web in all its contexts. Three main targets of trust are identified: content, services, and people. Trust in the content on the Web, including webpages, websites, and Semantic Web data is addressed first. Then, we move on to look at services including peer-to-peer environments and Web services. This includes a discussion of Web policy frameworks for access control. People are the final group, where we look at the role of trust in web-based social networks and algorithms for inferring trust relationships. Finally, we review applications that rely on trust and address how they utilize trust to improve functionality and interface.
Contents

1 Introduction 1
  1.1 Scope of Trust 3
  1.2 Trust in Content 4
  1.3 Trust in Services 4
  1.4 Trust in People 4

2 Trust in Content 7
  2.1 Trusting Web Pages 7
  2.2 The Semantic Web 10
  2.3 Trusting Semantic Web Information Sources 13
  2.4 Provenance on the Semantic Web 14
  2.5 Conclusions 17

3 Trust in Services 19
  3.1 Trust Management 20
  3.2 Computing Trust in Peer-to-Peer Systems 23
  3.3 Trust in Web-Services 26
  3.4 Conclusions 29

4 Trust in People 31
  4.1 Web-Based Social Networks 32
  4.2 Computing Trust in Social Networks 35
  4.3 Conclusions 42
5  Applications

5.1  Recommender Systems

5.2  News Syndication

5.3  Email Filtering

5.4  Peer to Peer Routing

6  Discussion and Conclusions

References
Almost since the inception of the web, trust has been a concern. The success of the web is based largely on its open, unmanaged nature; at the same time, that allows for a wide range of perspectives and intentions. It provides access to billions of web pages with staggering amounts of information; as a communication medium, the web connects people and services to one another for exchanging information and making transactions; some of the most exciting new activity on the web is social, with social networks and collaborative interaction. In all of these cases, there must be trust to foster successful interactions and to filter the abundance of information.

There are three major challenges to using trust on the web.

- Trust management: Jøsang et al. [28] define trust management as *The activity of creating systems and methods that allow relying parties to make assessments and decisions regarding the dependability of potential transactions involving risk, and that also allow players and system owners to increase and correctly represent the reliability of themselves and their systems*. More generally, trust management is the
process of determining who has access to what information or resources in a system, identity management, and delegation of trust. Essentially, instead of simply encrypting data for its protection, trust management establishes a set of policies and determines the credentials of a person or services to access the data [17]. Doing this accurately and efficiently in a variety of domains requires many approaches.

• Computing trust: The known trust relationships on the web are only a small fraction of the potential pairings. Furthermore, the number of pages, services, and users on the web is so large, that it is difficult to estimate how much trust there is between entities. For example, a user cannot possibly know how much to trust every other user and every page on the web. Instead, trust must be calculated from other available data. Depending on the context, the methods for doing that will vary.

• Applications using trust: Managing and computing trust are interesting problems, but ultimately they exist to provide trust information that can be used. Trust in people or content provides insight into how they should be treated within a system (e.g., should a person be given access to a resource or how much weight should a user give to some information). Building applications that take advantage of trust information and improve their functionality because of it requires an understanding of how trust relates to the system’s goals and how to integrate it. Doing this effectively is a challenge in all domains.

The proper way to address these challenges varies based on the context. For example, computing trust among web services via access control policies is quite different than computing trust between users in a social network. In this review, we consider trust in three domains: trust in content, trust in services, and trust in people. Once we have reviewed methods for managing and computing trust in those domains, we move on to applications. These integrate techniques from the domains to use trust for creating new functionality.
1.1 Scope of Trust

Trust has many meanings in computing, so we begin by describing the scope of this review with respect to the term. That is followed by brief descriptions of each section.

1.1 Scope of Trust

Within computer science, trust has been co-opted by many subfields to mean many different things. It is a descriptor of security and encryption [62]; a name for authentication methods or digital signatures [9]; a measure of the quality of a peer in P2P systems [96]; a factor in game theory [82]; a model for agent interactions [56]; a gauge of attack-resistance [104]; a component of ubiquitous computing [95]; a foundation for interactions in agent systems [13, 77]; and a motivation for online interaction and recommender systems [3]. On the web, many of these variations on trust are relevant. In a distributed, anonymous system like the web where services and information come from different sources, trust is especially important.

In this review, we treat trust as a concept that helps users (and agents) to make subjective decisions about content, services, or people when there is uncertainty. The breadth of these subjects excludes any single definition of “trust.” The subjective component, however, excludes cryptologic and many security issues from our scope.

Trust is largely a social concept, and its sociological and psychological attributes have been studied extensively. That work is largely relevant to the study of trust on the web, and it informs much of the research presented here. However, this review is scoped to focus on the science of trust on the web, and particularly computing with trust. We introduce algorithms, standards, and empirical studies as primary results, and social research only as it supports the computational work.

Trust has been an important topic in the agents community. While agents are often studied on the web, the research into trust and agents applies equally to non-web agents. This research is certainly applicable to many web contexts, but we have scoped this review to cover web trust only. Thus, agent-based trust is outside of what we cover in this review.
1.2 Trust in Content

The web is its content. It has revolutionized the way people access information, and the amount of information they have access to. It has done this by providing billions of pages on every conceivable topic, and tools like search engines have made it accessible. On top of pages, there are vast amounts of data stored in databases and XML formats. The success of the web is due largely to the fact that there is no centralized control of the web; anyone can say anything. At the same time, this lack of moderation raises the question of trust with respect to content. Instead of being able to make a simple trust decision about one central editor or moderator, the user has to make a series of decisions each time she accesses a page. Which pages and what data can be trusted? How is that trust established? How is information about its trust shared? Section 2 looks at questions of trust in content, from web pages to data on the Semantic Web.

1.3 Trust in Services

Automated services are an important part of the web. Peer-to-peer (P2P) systems and web services are both widely used and important. Trust is important in this context because sensitive information is often exchanged between services, and also because users rely on their successful completion. Since the interactions between these services is usually automated, the conditions for trust must be established ahead of time by the users.

In Section 3, we look at trust in P2P systems and web services. The main issues addressed are how to judge trust based on performance, how to propagate trust assessments in distributed environments, and how to specify policies that increase the trust in web services.

1.4 Trust in People

The web is a social environment, facilitating the exchange of ideas, documents, money, and goods. The social components are becoming increasingly visible. Social Networking is one of the largest movements on the web, with hundreds of millions of user accounts among hundreds
of different networks. Online communities supply a forum for discussion, ratings, and interaction. On the web, social trust and reputation are important factors that inform decisions about what to reveal to others and how to treat the information they provide. However, the web is also a very big place. The background information necessary for judging how much to trust an unknown person is often distributed and potentially private. Thus, methods for understanding, managing, computing, and applying trust are required.

Ultimately, users benefit from these social rankings because they can be used to judge things like the quality of information or the risk of a transaction. We can already see places where users have come to rely on trust and reputation, such as in eBay’s feedback or rating websites like Epinions. There is more that can be done with social trust, but it requires a better understanding of the properties and dynamics of the relationship. Trust is not a new concept in computing; it has been studied as a facet of security, encryption, authentication, and quality. Trust as a social relationship, however, has very different properties. Because they are social concepts, trust and reputation are fuzzier concepts than are normally treated by computer scientists. The social behavior of web users and the scale of web systems require new understanding and computational techniques. At the same time, the growth and evolution in the way the web is used demands solutions that rely on these advances.

The emergence of recent work to better understand the computational properties of social trust and reputation is timely and necessary. Researchers have been making progress on all fronts. We have developed new theories for managing and for understanding the properties of social trust and reputation relationships. That has laid the foundation for the many algorithms have recently been developed for computing trust relationships between people. Analysis of reputation systems have also led to results that help protect against deception. As this grounds for assessing trust and reputation has improved, a number of new applications have been developed that utilize trust. This brings the benefits of understanding the user’s social relationships into the applications that they already use.
References

References


References


References


Full text available at: http://dx.doi.org/10.1561/0400000006
References


References


