ERC Synergy Grant – White Paper

imPACT – Privacy, Accountability, Compliance, and Trust in Tomorrow's Internet

Principal Investigators (PIs) : Michael Backes, Saarland University
: Peter Druschel, Max Planck Institute for Software Systems
: Rupak Majumdar, Max Planck Institute for Software Systems
: Gerhard Weikum, Max Planck Institute for Informatics

Project summary

The Internet has evolved from a mere communication network used by a few million users to a global multimedia platform for communication, social networking, entertainment, education, trade, and political activism used by more than two billion users. While this transformation has brought tremendous benefits to society, it has also created new threats to privacy, safety, law enforcement, freedom of information, and freedom of speech. In today's Internet, users produce and consume information, potentially anonymously but desiring trust, and interact in complex ways that cannot be modeled adequately using existing techniques. Understanding this space of user interactions, and developing models and methods to ensure the continued viability of the future Internet is a daunting problem, yet one of utmost importance to society.

The *imPACT* project addresses the key challenges of providing **P**rivacy, **A**ccountability, **C**ompliance, and **T**rust (PACT) in tomorrow's Internet. We approach the problem using a cross-disciplinary and synergistic approach to understanding and mastering the different roles, interactions, and relationships of users and their joint effect on the four PACT properties. The focus is on principles and methodologies that are relevant to the needs of individual Internet users, have a strong potential to lead to practical solutions and address the fundamental long-term needs of the future Internet. We take on this challenge with a team of PIs consisting of international leaders in privacy and security, distributed systems, formal methods, program analysis and verification, and database systems, with input from outside experts in law, social sciences, economics, and business. By committing ourselves to this joint research, we are in a unique position to meet the grand challenge of realizing and maintaining the Internet's potential as an enabling platform for commerce, education, social exchange, democracy, and self-fulfillment.

1 Challenge

The Internet has evolved from a mere communication network used by tens of millions of users two decades ago, to a global multimedia platform for communication, social networking, entertainment, education, trade and political activism used by more than two billion users. Users have matured from consumers of information and services to content publishers, interactive participants, trade partners and sources of advice, opinion and expertise. An entire industry has developed that tracks and mines user activity and personal information for the purpose of targeted online advertising. Corporations, organizations and political parties alike have become keenly aware of the importance of understanding and influencing public opinion in and through the Internet.

While these trends have brought tremendous benefits to society, they have also created new threats to privacy, safety, law enforcement, freedom of information, and freedom of speech:

- When all users produce and consume, buy and sell, act and react, influence and be influenced, it is no longer feasible to distinguish between trusted publishers, merchants, and experts on the one hand, and untrusted consumers on the other. Users are tracked and profiled by providers; they may publish false or distorted content, manipulate other users, and steal or vandalize their content. Even worse, users and providers can act under false or multiple online identities, making it difficult to hold them accountable for their actions.

- In today's Internet, data is increasingly stored and computation executed on third-party Cloud platforms; while third-party applications are executed in client-side environments like browsers and smart phones. These techniques provide ubiquitous access, but also introduce new threats to privacy, trust, and compliance.

Understanding this space of competing phenomena and developing models and methods to ensure the continued viability of the future Internet is a daunting problem of utmost importance to society.

The imPACT project addresses the following key challenges:

1. Protecting users' *privacy* when publishing content on the Internet and when participating in online communities;

2. Ensuring *accountability* of users and providers so that misbehavior can be detected and the culprits identified;

3. Ensuring the *compliance* of software and services with user expectations, applicable laws and provider policies; and

4. Assessing the *trustworthiness* of information and services that a user consumes, while maintaining free choice among diverse sources and providers.

The resulting properties are relative to user and provider "specifications" of their desired policies. We refer to the four properties as PACT, for privacy, accountability, compliance, and trust.

2 Threats in Today's Internet

Numerous known risks and frequent incidents illustrate that today's Internet fails to ensure the PACT properties.

Online *privacy* remains a challenge for users, providers, and legislators alike. Users tend to reveal personal information without considering the wide circulation, easy accessibility and permanent nature of online data. Many cases reported in the press show the resulting risks, which range from public embarrassment to disadvantages when applying for jobs or insurance, to personal safety and property risks when stalkers, sexual offenders, or burglars learn users' whereabouts online. Legislators have responded by tightening privacy regulations (e.g., mandating user control of their personal information), but providers arguably lack the technology to fully comply with such regulations, and users lack models and tools to reason about their privacy choices effectively.

Online *accountability* presents another serious challenge. Today, users can assume multiple false online identities to commit fraud in auction sites like eBay; slander or bully other users in online social networks, blogs or chat rooms; distribute or acquire illegal content; manipulate recommendation systems and public opinion; or commit data theft, identity theft, or blackmail. Cases like these occur in the Internet on a daily basis. Many legislators have responded, for instance, by requiring user registration at all Internet access point. However, in the current Internet, it does not take much technical sophistication for a user to obscure the actual source of their online activity. Moreover, technology to address this problem must accommodate users with a legitimate need for strong anonymity, such as whistle blowers or political dissidents in countries that don't respect free speech.

Regarding *compliance*, providers currently lack technology to guarantee that their services fully comply with applicable laws, regulations, and their own policies; nor do they have the means to prove their compliance to regulators and users. As a result, cases of demonstrated (e.g., Google street view's collection of users' wifi traffic) and suspected (Facebook's alleged non-compliance with EU regulations) non-compliance abound. Technology to ensure and demonstrate compliance is complicated by providers' legitimate need to protect their trade secrets.

Finally, *trust* in online information and services is complicated by the fact that more information is contributed by ordinary users and crowdsourced. For instance, false statements about prominent individuals, such as celebrities and politicans, may be found in the Internet more frequently than the actual facts. Moreover, current search algorithms that only consider co-occurrence of terms rather than trustworthiness of information can propagate and perpetuate such falsehoods by returning them in search results simply because many websites contain the false information. Ensuring trust in online information and services requires technology to track the provenance of online information, and the trustworthiness of information sources.

3 Goal: Ensuring PACT in the Future Internet

The imPACT project embarks on an expedition to develop fundamental principles and methodologies to enable an Internet that respects user's privacy, holds users and providers accountable for their actions, ensures compliance with user and provider policies, and ensures trust in services and information. In contrast to existing security and privacy technology, a paradigm shift is needed to cope with the wealth of user-to-user and user-to-provider interactions; the need for privacy, freedom of speech and freedom of information; and the needs of the Internet economy and law enforcement. It is important to understand that many of the PACT challenges affecting today's Internet cannot be fully addressed by technology alone. Our focus will be on developing and evaluating novel principles, models, tools and technology that are relevant to Internet users, that have a strong potential to lead to practical solutions, and that address the fundamental long-term needs of the future Internet within appropriate legal, regulatory, and economic frameworks.

This research strives for compatibility with existing business models, and includes exploring business-to-consumer issues as well as opportunities for deploying novel PACT solutions in field trials with Internet services.

4 Approach and Synergies

The central theme of the project are the PACT properties, which are key to maintaining the viability of the future Internet. The four properties have *separately* received some attention in prior work: privacy and compliance to a considerable, accountability and trust to a lesser, extent. Existing models for security and privacy tend to be focused on user authentication, authorization, and access control. These concepts are too narrow to reflect the complex interrelationships among users, providers, and services in today's Internet (Web 2.0). The true challenge lies in understanding and mastering the different roles, interactions, and relationships of users and their *joint* effect on the four PACT properties. This is precisely the aim of this project.

The challenge will be addressed by a team of researchers from relevant subdisciplines within computer science, and will be informed by outside experts in law, social sciences, and business. The team of PIs consists of international leaders in privacy and security (Backes), distributed systems (Druschel), formal methods, program analysis and verification (Majumdar), database systems and knowledge management (Weikum). All four PIs have an extensive track record related to one of the PACT properties.

Synergies among these research areas are critical to the success of the project. Naturally, security and privacy expertise lies at the heart of the effort. Ensuring that methods and solutions are practical, scalable, usable and deployable requires the expertise of experimental systems researchers. Developing appropriate ways of specifying policies and requirements, and ensuring that the trusted computing base complies with these policies and requirements requires formal methods and verification researchers. Understanding the implications of large-scale data aggregation and mining for privacy and information trust requires the expertise of database systems researchers. In addition to these core competencies, our team has complementary background in methodologies like cryptographic methods, correctness by design and verification, building distributed systems and distributed testbeds, and Web-scale data analysis. *By teaming up and committing ourselves to this joint research, we are in a unique position to meet the grand challenge of unifying the PACT dimensions and laying a new foundation for their holistic treatment.*

5 Scope and Potential Impact

Our approach toward these goals is a long-term strategy with emphasis on fundamental models and methods that will provide a foundation of PACT for current and future Internet applications. The foundational work will be accompanied by software prototyping and field-trial experiments for testing of building blocks, demonstration of benefits, and reality checks. We focus on software solutions and on user-centric solutions and business-to-consumer (B2C) settings. While the project's focus is on information technology and computer science, we will consult with scholars and experts in law, social science, economics and business.

Today's Internet is in a deep crisis of confidence. The high cost of cybercrime (US$ 114B annually according to a recent estimate by Symantek) compounds with the hard to quantify cost of privacy loss and lost opportunity due to security concerns. Users, organizations and even legislators are torn between extremes of paranoia (e.g., the Internet is an unsafe place that must be tightly regulated), or cynicism and naïveté (e.g., privacy is a thing of the past; all digital content should be free). To realize and maintain the Internet's vast potential as an enabler of commerce, education, social exchange, self-fulfillment, democracy and free speech, a major effort is required. *Countering the current trends, restoring confidence and ensuring the PACT properties in the Internet is of fundamental importance to society*. The imPACT project takes on this challenge.