Information Flow Experiments (Extended Abstract)

Michael Carl Tschantz
mct@berkeley.edu
UC Berkeley

Amit Datta
amitdatta@cmu.edu
Carnegie Mellon University

Anupam Datta
danupam@cmu.edu
Carnegie Mellon University

Jeannette M. Wing
wing@microsoft.com
Microsoft Research

Web Data Usage Detection. Concerns about privacy have led to much interest in determining how third-party associates of first-party websites use information they collect about the visitors to the first-party website. Mayer and Mitchell provide a recent presentation of research that tries to determine what information these third-parties collect [6]. Others have attempted to determine what these third-parties *do* with the information they collect [1], [5], [16], [21]. We call this problem *web data usage detection* (WDUD).

The researchers involved in WDUD each propose and use various analyses to determine what information is tracked and how it is used. They primarily design their analyses by intuition and do not formally present or study their analyses. Thus, questions remain: (1) Are the analyses used correct? (2) Are they related to more formal prior work?

To answer these questions, we must start with a formal framework that can express the problem and the analyses. In essence, each of these works is conducting an information flow analysis: the researchers want to know when information flows to a third-party and where it goes from there. Thus, the natural starting point for such a formalism is prior research on *information flow analysis* (IFA). However, despite the great deal of research on IFA (see [13] for a survey), we know of no attempt to relate or inform WDUD research with the models or techniques of IFA, even in an informal manner.

We believe this disconnect exists since, unlike traditional IFA, the analyst has no access to the program running the third-party service, little control over its inputs, and a limited view of its behavior. Thus, the analyst does not have the information presupposed by traditional IFAs. To understand the WDUD problem as an instance of IFA requires a fresh perspective on IFA.

Other "Hidden" IFA Problems. The implicit assumptions underlying much of IFA research also obscure its connection to other areas of research.

For example, the cryptography community has much work on identifying illicit flows of files. Such work has included *watermarking* [15], [20], in which a key that links to the identity of the person to whom the publisher sold the copy is embedded in the work.

Closely related is the detection of plagiarism. One approach the publisher can use for this problem is to employ a *copyright trap*: deliberately unusual (typically, false) information inserted into reference works to detect copying [8].

Organizations handling sensitive data are concerned about data misuse. For example, governments are concerned with employees leaking classified documents to reporters or foreign spies. For ethical reasons and to comply with regulations, such as the HIPAA Privacy Rule [9], healthcare providers limit the use of personal health information. Thus, organizations have adopted a variety of methods to discourage the misuse of such data by their employees [12], [17]. For example, investigators have employed *Barium meals*, a watermarking-like analysis [22].

In essence, these works are all IFAs. In particular, the analyst, who is aligned with the copyright holder or organization, would like to determine whether a system (typically a personal computer or person) is enabling an illicit flow of information. However, those working on these problems have not typically discussed them as such since they do not fit into the traditional IFA setting. In particular, the analyst has little if any access or control over the analyzed system. Like with WDUD, the analyst must investigate an uncontrolled black box. Indeed, we find that some of the intuitive approaches used in WDUD are related to cryptographic measures used in piracy detection.

Goal. Our goal is to systematize the information flow problems and analyses common to these areas of research. To do so, we identify the limited abilities of the analyst in these problems. as a form of analysis between the extremes of white box program analysis and black box monitoring. We show that the ability of the analyst to control some inputs during an investigation enables *information flow experiments* that manipulate the system in question to discover its use of information without a white box model of the system. Our framework provides

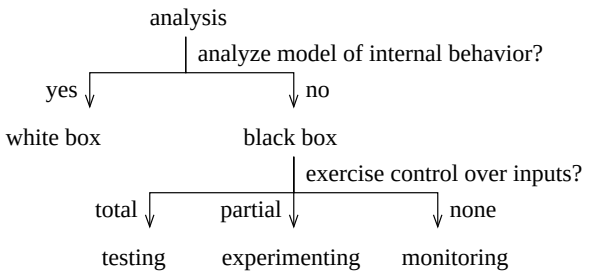

Fig. 1. Taxonomy of analyses

a fresh perspective both on our diverse set of motivating applications and on IFA by allowing us to elucidate and challenge approaches in these areas and in IFA.

The overarching contribution of this work is relating IFA in these nontraditional settings to experiments designed to determine causation. To do so, we prove a connection between information flow and causality, which allows us to reduce these problems to well understood empirical ones. In particular, it allows us to use statistical analyses in the place of traditional methods of IFA, such as program analysis.

Overview. Goguen and Meseguer introduced *noninterference* to formalize when a sensitive input to a system with multiple users is protected from untrusted users of that system [3]. Intuitively, noninterference requires that the system behaves identically from the perspective of untrusted users regardless of any sensitive inputs to the system.

Information flow analysis (IFA) is a set of techniques to determine whether a system has noninterference (or similar properties). Most IFA methods, such as type systems [13], [18], are inappropriate for WDUD since they require *white box* access to the program. That is, the analyst must be able to study and/or modify the code. In our applications, the analyst must treat the program as a *black box*. That is, the analyst can only study the I/O behavior of the program and not its internal structure. Black box analyses vary based on how much access they require to the system in question. Figure 1 shows a taxonomy of analyses.

In our motivating problems, the analyst

1) has no model of or access to the program running the system,
2) cannot observe the internal states of the system,
3) has limited control over and knowledge of the environment of the system,
4) can observe a subset of the system's outputs, and
5) has control over a subset of the inputs to the system.

TABLE I. EXPERIMENTAL SCIENCE, IFA, AND WDUD COMPARED

Experimental Science	Information Flow	WDUD
natural process	system in question	Google etc.
population of units	subset of interactions	browser instances
factors	input channels	visitor behavior
treatments	controlled inputs	behavior profiles
noise factors	uncontrolled channels	advertisers, others
response variables	observed output channels	sequences of ads
effect	interference	use of data

We will call performing IFA in this setting *experimenting*. Experiments may be viewed as an interactive extension of a limited form of execution monitoring that allows for analyst inputs but limits the analyst to only observing a subset of system I/O. Unfortunately, while analyses exist for the case where the analyst has total control over inputs (testing) (e.g., [23]), we know of none for our case of experimenting with limited control over input.

Prior work shows that no monitor can detect information flows [7], [14], [19]. We argue that experiments, with the additional ability to control some inputs to the system, do not improve upon this situation. In particular, we prove that no non-degenerate analysis can be sound for interference or for noninterference, even on deterministic systems.

However, we also show a connection between noninterference and causality: noninterference corresponds to a lack of an effect. This result allows us to repose WDUD as a problem of statistical inference from experimental data using causal reasoning. In particular, we use Pearl's formalization of *effect* using *structural equation models* (SEMs) [10]. We show that an SEM model of a system has a causal effect if and only if an automaton model of the system has interference.

With this connection in hand, we propose to approach our problems as empirical studies of causality employing experimental designs from the natural sciences. We summarize the relationship in Table I.

After designing and running an experiment, scientists must analyze the data collected. In particular, they must quantify the probability that the collected responses could have occurred by chance through an unlucky random assignment of units to treatments. A common approach to quantifying experimental results is by *significance testing* [2]. The possibility of an unlucky assignment of units is formalized as a *null hypothesis* that states that the groups differ by chance. A *statistical test* of the data provides a *p-value*, the probability of seeing results at least as extreme as the observed data under the assumption that the null hypothesis is true. A small p-value implies that the data is unlikely under

the null hypothesis. Typically, scientists are comfortable rejecting the null hypothesis if the p-value is below a threshold of 0.05 or 0.01 depending on field. Rejecting the null hypothesis makes the alternative hypotheses more plausible. In our case, the null hypothesis is that the system in question has noninterference and the alternative of interest is the system has interference.

We recommend *non-parametric* tests, which do not require assuming a family of distributions and instead treat the generating distribution as a black box. In particular, we will focus on *permutation tests* (see e.g., [4]). Crucially, permutation tests (also known as *randomization tests*) allow cross-unit interactions [11], which can occur in WDUD studies.

We show that permutation tests perform well in practice by using them in our own WDUD experiment, which systematizes prior studies [1], [5], [16], [21].

Contributions. Our methodology is supported by an unbroken chain of contributions:

1) a systematization of nontraditional IFA,
2) a proof of connection between IFA and causality,
3) an experimental design leveraging this connection,
4) a statistical approach to analyzing experimental data, and
5) a systematization prior studies under this unified method.

These contributions are each necessary for creating a chain of sound reasoning from intuition about vague problems to rigorous quantified results in a formal model. This chain of reasoning provides a systematic, unifying, view of these problems, which leads to a concrete methodology based on well studied scientific methods. While the notion of experimental science is hardly new, our careful justification provides guidance on the choices involved in actually conducting an information flow experiment.

We further support our reasoning with our own experiments to illustrate the abstract concepts we present. These results may also be of independent interest to the reader, but for reasons of space, they are not presented in this abstract.

The systematization of experimental approaches to security is becoming increasingly important as technology trends (e.g., Cloud and Web services) result in analysts having limited access to and control over systems whose properties they are expected to study. This paper provides a useful starting point towards such a systematization by providing a common model and a shared vocabulary of concepts that ties together seemingly disparate areas of security and privacy by placing them in the context of causality, experimentation, and statistical analysis.

REFERENCES

[1] BALEBAKO, R., LEON, P., SHAY, R., UR, B., WANG, Y., AND CRANOR, L. Measuring the effectiveness of privacy tools for limiting behavioral advertising. In *Web 2.0 Security and Privacy Wksp.* (2012).

[2] FISHER, R. A. *The Design of Experiments*. Oliver & Boyd, 1935.

[3] GOGUEN, J. A., AND MESEGUER, J. Security policies and security models. In *IEEE Symp. on Security and Privacy* (1982), pp. 11–20.

[4] GOOD, P. *Permutation, Parametric and Bootstrap Tests of Hypotheses*. Springer, 2005.

[5] GUHA, S., CHENG, B., AND FRANCIS, P. Challenges in measuring online advertising systems. In *10th ACM SIGCOMM Conf. on Internet Measurement* (2010), pp. 81–87.

[6] MAYER, J. R., AND MITCHELL, J. C. Third-party web tracking: Policy and technology. In *IEEE Symp. on Security and Privacy* (2012), pp. 413–427.

[7] MCLEAN, J. A general theory of composition for trace sets closed under selective interleaving functions. In *1994 IEEE Symp. on Security and Privacy* (1994), p. 79.

[8] MONMONIER, M., AND DE BLIJ, H. J. *How to Lie with Maps*, 2 ed. University of Chicago Press, 1996.

[9] OFFICE FOR CIVIL RIGHTS. Summary of the HIPAA privacy rule. OCR Privacy Brief, U.S. Department of Health and Human Services, 2003.

[10] PEARL, J. *Causality*, second ed. Cambridge University Press, 2009.

[11] ROSENBAUM, P. R. Interference between units in randomized experiments. *J. the American Statistical Association 102*, 477 (2007), 191–200.

[12] RSA LABS. RSA data loss prevention.

[13] SABELFELD, A., AND MYERS, A. C. Language-based information-flow security. *IEEE Journal on Selected Areas in Communications 21*, 1 (2003), 5–19.

[14] SCHNEIDER, F. B. Enforceable security policies. *ACM Trans. Inf. Syst. Secur. 3*, 1 (2000), 30–50.

[15] SWANSON, M., KOBAYASHI, M., AND TEWFIK, A. Multimedia data-embedding and watermarking technologies. *IEEE 86*, 6 (1998), 1064–1087.

[16] SWEENEY, L. Discrimination in online ad delivery. *Commun. ACM 56*, 5 (2013), 44–54.

[17] SYMANTEC. Symantec data loss prevention.

[18] VOLPANO, D., IRVINE, C., AND SMITH, G. A sound type system for secure flow analysis. *J. Comput. Secur. 4*, 2-3 (1996), 167–187.

[19] VOLPANO, D. M. Safety versus secrecy. In *6th Intl. Symp. on Static Analysis* (1999), Springer-Verlag, pp. 303–311.

[20] WAGNER, N. R. Fingerprinting. In *1983 IEEE Symp. on Security and Privacy* (1983), p. 18.

[21] WILLS, C. E., AND TATAR, C. Understanding what they do with what they know. In *2012 ACM Wksp. on Privacy in the Electronic Society* (2012), pp. 13–18.

[22] WRIGHT, P. *Spycatcher: The Candid Autobiography of a Senior Intelligence Officer*. Viking Adult, 1987.

[23] YUMEREFENDI, A. R., MICKLE, B., AND COX, L. P. Tightlip: keeping applications from spilling the beans. In *4th USENIX Conf. on Networked Systems Design and Implementation* (2007), pp. 12–12.