

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent of: Roy Shkedi Attorney Docket No.: 58717-0005IP1
U.S. Patent No.: 8,494,904
Issue Date: July 23, 2013
Application No.: 13/585,779
Filing Date: August 14, 2012
Title: METHOD AND STORED PROGRAM FOR
ACCUMULATING DESCRIPTIVE PROFILE DATA
ALONG WITH SOURCE INFORMATION FOR USE IN
TARGETING THIRD-PARTY ADVERTISEMENTS

DECLARATION OF DR. MICHAEL D. MITZENMACHER

I declare that all statements made herein on my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable under Section 1001 of Title 18 of the United States Code.

Date: 6/13/25

By: Michael Mitzenmacher
Dr. Michael D. Mitzenmacher

TABLE OF CONTENTS

I.	Introduction.....	4
II.	Qualifications and Background Information	4
III.	Combined Teachings of Robinson and Jaye (Robinson-Jaye).....	7
	A. Overview of Robinson and Jaye	7
	1. Robinson Discloses a Central System That Tracks, in User Profiles, Web Activity of Users as they Visit Tracking-Enabled Web Pages.	7
	2. Jaye Discloses a Network of Local Servers and an Enterprise Server in Which the Local Servers Collect Local User Profiles That the Enterprise Server Compiles in Global User Profiles.	11
	B. Robinson-Jaye Combination.....	16
	1. A POSITA Would Have Been Led to Combine Robinson’s and Jaye’s Teachings.....	16
	2. The Combined Robinson and Jaye System.....	24
	3. In the Robinson-Jaye Combination, a Global User Profile Stored in the Central System is Constructed from Local User Profiles Stored on Multiple Different Host Machines.....	38
	C. An Example of an Implementation of Robinson-Jaye Combination	42
	1. Network Configuration	42
	2. Registration of Local User Identifier in Global User Profile.....	45
	3. Upload of Local Profile Information to Global User Profile.....	52
	4. Storage of Local User Identifiers and Local Profile Information from Multiple Host Machines and for Multiple Users in Global User Profiles	55
	5. Targeting Advertising for Users.....	58
IV.	Combined Teachings of Robinson, Jaye, and Coleman (Robinson-Jaye-Coleman).....	66
	A. Coleman Discloses Gathering Information about Users to Create User Profiles, Verifying the Gathered Information, and Selling the Information to Advertisers for Targeted Advertising.....	67
	B. A POSITA Would Have Been Led to Combine Robinson’s, Jaye’s, and Coleman’s Teachings.....	71
	C. An Example of an Implementation of Robinson-Jaye-Coleman.....	79
V.	Manner in Which The Prior Art References Render the ’904 Patent Claims Obvious.....	82

A.	[GROUND 1A] Robinson-Jaye Renders Obvious Claims 1, 3-7, 9-11, 13-17, 19-21, 23-27, and 29-30.	83
B.	[GROUND 1B] Robinson-Jaye-Coleman Renders Obvious Claims 7-9, 17-19, and 27-29.	98
VI.	Level of Ordinary Skill in the Art	100
VII.	Legal Principles	102
A.	Terminology.....	102
B.	Legal Standards.....	102
C.	Anticipation.....	103
D.	Obviousness	103
VIII.	Conclusion	108

I. INTRODUCTION

1. I have been retained on behalf of LiveIntent, Inc. (LiveIntent) to offer technical opinions related to U.S. Patent No. 8,494,904 ('904 Patent) (LIVEINTENT1001). I understand that LiveIntent is requesting that the Patent Trial and Appeal Board (Board) to institute an *inter partes* review (IPR) proceeding of the '904 Patent. I have been asked to provide my independent analysis of the '904 Patent based on the prior art publications and other materials cited in this declaration and based on my own knowledge, education, experience, and expertise in the relevant fields.

II. QUALIFICATIONS AND BACKGROUND INFORMATION

2. My qualifications are summarized here and explained in more detail in my curriculum vitae, which is attached as Appendix A to this report.

3. I am currently employed as a Professor of Computer Science at Harvard University. Specifically, I am the Thomas J. Watson, Sr. Professor of Computer Science in the School of Engineering and Applied Sciences. I joined the faculty of Harvard as an Assistant Professor in January 1999. I was promoted to Associate Professor in 2002 and to Professor in 2005. In 2010, I began a three-year term as Area Dean, which is essentially equivalent to what other schools call Department Chair, of Computer Science, and held that position through June 2013.

I also served as the Area Co-Chair for Computer Science from July 1, 2018, to June 30, 2019. My work address is 150 Western Ave, Allston, MA 02134.

4. I received my undergraduate degree in Mathematics and Computer Science from Harvard College in 1991. I received a Certificate of Advanced Study in Mathematics from Cambridge University in 1992. I received a Ph.D. in Computer Science from the University of California at Berkeley in 1996. From August 1996 to January 1999, I was employed as a Research Scientist at Digital Systems Research Center.

5. I have published over 250 research papers in computer science conferences and journals. My general area of research includes algorithms and data structures, as well as the analysis of probabilistic processes. I am listed as an inventor or co-inventor on 19 issued patents and am the co-author of a textbook, entitled *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, published by Cambridge University Press. I am a Fellow of the Institute of Electrical and Electronics Engineers (IEEE) and a Fellow of the Association for Computing Machinery (ACM). In addition, I have consulted for several companies while I have been a professor, such as Microsoft, Mitsubishi, Akamai, ITA software, Fluent Mobile, Adverplex, and e-Harmony.

6. I teach courses on various computer science topics to undergraduate and graduate students at Harvard. Among other topics, I have regularly taught a

class on algorithms and data structures for information, networking, and communication to Harvard undergraduate and graduate students.

7. I have both done research in the area of electronic commerce, as well as worked for companies related to electronic commerce. For example, I have written papers regarding how the use of Groupon can affect ratings (such as Yelp ratings) for businesses, on the issue of information asymmetry in pay-per-bid auction sites, and on algorithms for finding statistically significant frequent item sets. Also, as examples, I have served on the scientific advisory board for e-harmony (a dating service with millions of users), and worked at companies such as Google, ITA software (which was bought by Google and serves as the basis for Google flights) and Adverplex/Cogolabs (which works in the space of web and Internet advertising).

8. Much of my work involves issues relating to networking and communications. I regularly serve on program committees for conferences in networking, algorithms, and communication. For example, I have served on the program committee multiple times for the SIGCOMM conference, which is described on the conference homepage as follows: “SIGCOMM is the flagship annual conference of the ACM Special Interest Group on Data Communication (SIGCOMM) on the applications, technologies, architectures, and protocols for computer communication.” <https://sigcomm.org/events/sigcomm-conference>.

Similarly, I have served several times on the Program Committee for NSDI, the USENIX Symposium on Networked Systems Design and Implementation. I have written papers on networking that have been published in the IEEE/ACM Transactions on Computer Networking, the SIGCOMM conference, the INFOCOM conference, and other major venues for networking research. My graduate course entitled “Algorithms at the end of the wire” covers many subjects at the intersection of networking and algorithms.

III. COMBINED TEACHINGS OF ROBINSON AND JAYE (ROBINSON-JAYE)

9. I provide below an analysis of U.S. Patent No. 5,918,014A (Robinson) (LIVEINTENT1004) and U.S. Patent No. 6,415,322B1 (Jaye) (LIVEINTENT1005) and my opinion of what I believe a person of ordinary skill in the art having the level of knowledge and skill as of the invention date of the '904 Patent (POSITA) (discussed in greater detail in Section VI below) would have found obvious in view of the combined teachings of Robinson and Jaye (Robinson-Jaye).

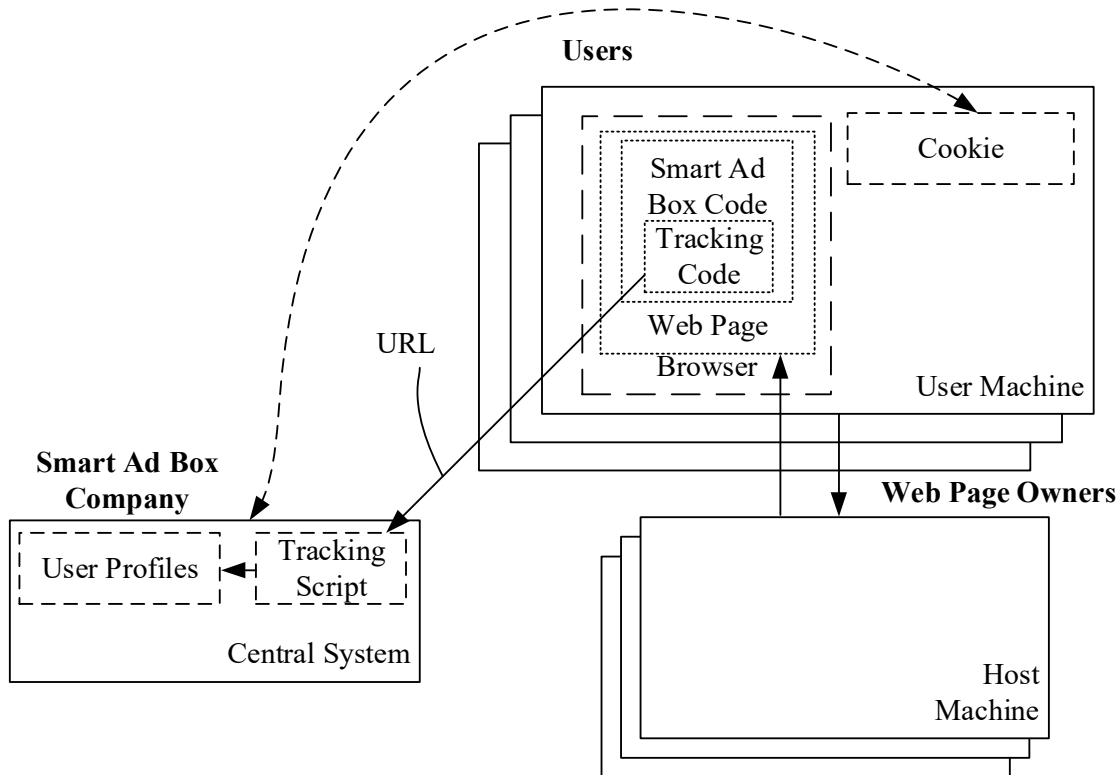
A. Overview of Robinson and Jaye

1. Robinson Discloses a Central System That Tracks, in User Profiles, Web Activity of Users as they Visit Tracking-Enabled Web Pages.

10. As illustrated in the diagram below of Robinson's system, Robinson describes a company providing Smart Ad Box services and running a central

system including a central database and server that stores tracking information related to users as they operate user machines to access tracking-enabled web pages and sites on host machines. LIVEINTENT1004, 3:62-64, 8:40-60, 9:4-45, 10:23-29, 13:43-49, 13:61-64; *see* LIVEINTENT1004, 2:49-51, 5:10-6:51, 7:66-8:20, 13:22-14:67. Smart Ad Boxes facilitate the display by tracking-enabled pages of web advertising, selected using “[s]pecial software algorithms.” LIVEINTENT1004, 4:8-13, 8:64-9:7; *see* LIVEINTENT1004, 4:39-4:67, 15:14-19:26. To provide the Smart Ad Box, code (e.g., HTML code) on the page draws a graphical box to accommodate an advertisement and updates the central database with information collected about the user. LIVEINTENT1004, 4:39-67, 8:64-9:35, 9:65-10:7, 13:24-30.

Petitioner-Provided Diagram 1: Robinson's System



11. The central system tracks user activities to generate user profiles. LIVEINTENT1004, 3:61-4:6; *see* LIVEINTENT1004, 15:43-54, 18:60-67. For example, as shown in the diagram above, each time a user references a tracking-enabled page, Smart Ad Box code on the page calls on a tracking script that: (1) is accessible through one URL, (2) is hosted on the central system, and (3) stores or accesses a cookie on a user machine to identify the user from session to session. LIVEINTENT1004, 9:12-16, 9:21-35, 9:36-53, 10:8-14. With the tracking script, the central system collects information relating to the user and user's visit to a

tracking-enabled page—such as a web site the user visited, a frequency of such visits, a nature of content of the web site visited by the user, an identifier of the page (e.g., one of several web pages on the web site) the user visited, the length of time spent on the page, the cookie identifying the user, the amount of money spent by the user while visiting the page, and other information related to the user—and aggregates this information in a profile of the user that the central system uses to identify similarities with profiles of other users. LIVEINTENT1004, abstract, 2:31-48, 9:8-35, 10:8-14, 15:16-54, 18:60-19:5; *see, e.g.*, LIVEINTENT1004, 5:30-36, 5:41-45, 6:38-42, 9:25-27, 13:65-14:8 (distinguishing pages and sites hosting those pages).

12. The information in the central system is used to determine an advertisement to present to the user, e.g., in the Smart Ad Box. LIVEINTENT1004, 15:15-62; *see* LIVEINTENT1004, 2:23-3:15, 3:64-4:6, 4:51-64, 16:49-18:67, fig.1. For instance, Robinson describes selecting an advertisement to display to the user based on a degree of similarity or closeness between that user and other users (e.g., based on whether the user and the other users have visited similar or the same pages). LIVEINTENT1004, 15:15-21, 15:43-50; *see* LIVEINTENT1004, 2:58-3:15, 3:65-4:6, 16:1-18:59, fig.1. In one particular example, Robinson describes predicting a user's interest in an advertisement based on tracked information about the user and the other users. LIVEINTENT1004,

15:15-19, 16:3-38, 16:53-67, 17:22-28, 18:18-21, 18:55-59; *see*

LIVEINTENT1004, 2:63-67, 16:1-18:59.

2. Jaye Discloses a Network of Local Servers and an Enterprise Server in Which the Local Servers Collect Local User Profiles That the Enterprise Server Compiles in Global User Profiles.

13. Jaye teaches a “distributed identification scheme” that “allows individual [local] servers to [simultaneously] control their own local identification scheme” for collecting local profile information *and* “collaborate with other servers” to share information about users that would be useful for tailoring web pages to users visiting web pages. LIVEINTENT1005, 1:62-2:11; *see* LIVEINTENT1005, 1:18-61.

14. As Jaye describes, web pages (e.g., hosted by local servers) can collect information about users (e.g., using client machines) as the users visit these web pages. LIVEINTENT1005, 1:18-34; *see* LIVEINTENT1005, 1:39-61. Owners of these web pages can use this collected information to “customize [the] Web pages on the fly for each user so that a user that initially selects particular pages is presented with the opportunity to select more of the type of pages in which that user is expected to be interested.” LIVEINTENT1005, 1:34-38; *see* LIVEINTENT1005, 1:39-61. A POSITA therefore would have recognized the benefits of enabling local servers or host machines to customize web pages “on the

fly” for each user based on a local user profile created from the user’s locally tracked activity on the host machines. LIVEINTENT1005 at 1:28-45.

15. But information sharing between the local servers, as Jaye observes, is “extremely difficult” to achieve in a way that preserves user privacy.

LIVEINTENT1005, 1:40-61. A single, shared global identifier for a user, as Jaye explains, might violate the user’s privacy because a local server would still be able to access information tied to the shared global identifier even if the local server chooses to stop collaborating and sharing information with other local servers. LIVEINTENT1005, 1:45-61.

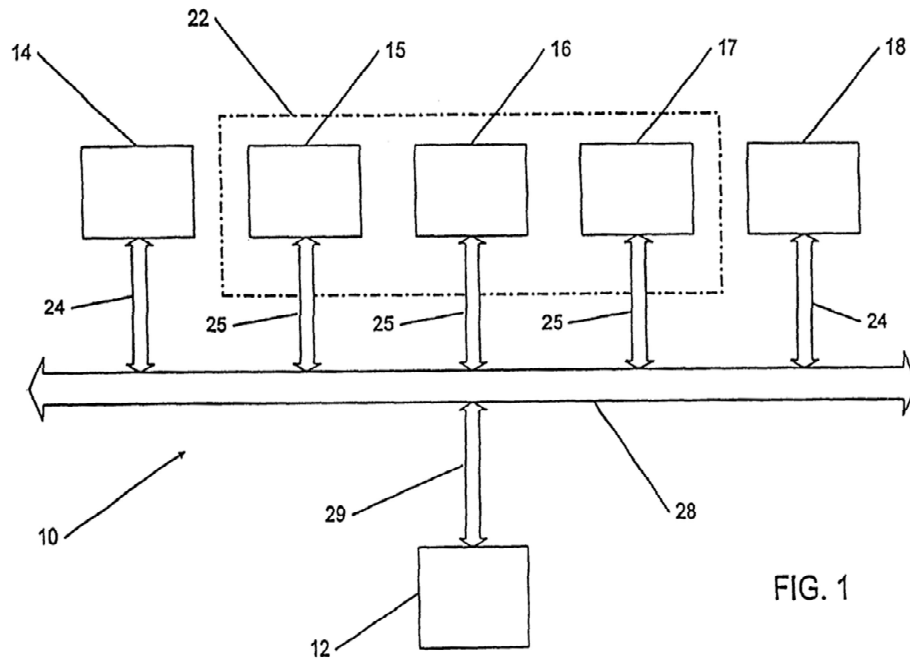
16. This is why Jaye’s system limits access and maintenance of the user’s global identifier to only the central system. *See* LIVEINTENT1005, 1:39-2:28. Specifically, Jaye teaches a “distributed identification scheme.”

LIVEINTENT1005, 1:62-2:11. This distributed identification scheme allows local servers “to control their own local identification scheme and to collaborate with other local servers or domains at the discretion of an enterprise server,” which uses its own “global user identifier” to identify a user who interact with the local servers without revealing an identity of the user to interested outside parties, such as advertisers. LIVEINTENT1005, 2:2-10. The local servers maintain their own local user identification schemes such that a local server’s local identifier for a particular user “may be hidden from other local servers” but is linked to a global user

identifier maintained by the enterprise server—“known only to the enterprise server” but “kept secret from the local servers.” LIVEINTENT1005, 2:15-31. Such local identification schemes by local servers, as Jaye teaches, would allow the local servers to identify users and obtain information about the users over multiple sessions. LIVEINTENT1005, 1:45-61, 2:1-6, 5:20-51, 8:53-55. And the enterprise server can anonymously compile a “global user profile” for a user from the “local user profiles generated by the local servers.” LIVEINTENT1005, 2:11-13.

17. Referring to Figure 1 of Jaye below, Jaye discloses an example of a computer network implementing a distributed identification scheme that includes local servers (e.g., local server 15 or 17), a client 12 (e.g., a user’s machine), and an enterprise server 16. LIVEINTENT1005, 4:5-30. To facilitate building of local user profiles concerning the client 12, the local servers 15/17 assign their own local user identifiers to the client 12, and Jaye describes this process through reference to local cookies deposited by the local servers onto the user machine client(s). LIVEINTENT1005, 2:2-6, 2:18-21, 2:55-58, 3:3-5, 3:14-17, 3:48-49, 5:49-51, 5:55-67, 6:67-7:11, figs.2, 4; *see* LIVEINTENT1005, 5:31-6:11, 6:12-7:26, 8:40-10:41. Similarly, Jaye describes a global user identifier being assigned by the enterprise server 16 to the client (e.g., a global cookie similar to the cookie described in both Robinson), and the global user identifier being linked to the local user identifiers and used to build a global user profile using the local user profiles.

LIVEINTENT1005, 2:6-24, 5:51-56, 6:6-11, 8:10-52; *see* LIVEINTENT1005, abstract, 2:50-3:39, 3:49-55, 5:31-6:11, 7:37-8:39, 10:22-41, 10:42-13:41, 14:22-25, figs.2-4.



LIVEINTENT1005 fig.1.

18. Jaye also describes an example of a process by which the enterprise server 16 might assign a global user identifier to the user and link the global user identifier to a local user identifier. Specifically, in its example, Jaye details how the local server may be configured to use a forced transfer or URL redirection from a web page of the local server to the enterprise server 16. LIVEINTENT1005, 6:32-46, 7:46-8:39, fig.2 (reference numeral 36), fig.3. When the client 12 requests

a web page from the local server, the local server initiates registration of a local user identifier by setting the identifier and then redirecting (e.g., forcing a transfer of) the client 12 to the enterprise server 16; Jaye offers an example of redirecting by inserting on the web page a special URL. LIVEINTENT1005, 6:32-38; *see* LIVEINTENT1005, 2:43-48, 4:64-5:4, 6:32-46, fig.2 (reference numeral 36); *see also* LIVEINTENT1005, 6:47-54 (Jaye providing another example of redirection in which the user is transferred to the enterprise server via an HTTP response message that causes the user's browser to look for a different URL).

19. The special URL points the client 12 to the enterprise server 16 and provides a graphics image having zero width and height akin to what is known in the art as a "pixel tag." LIVEINTENT1005, 6:38-6:46; *see* LIVEINTENT1005, 2:43-48, 4:64-5:4, 6:32-46. The local server appends information to the end of the special URL—including a local server identifier and client information such as the local user identifier—that the enterprise server 16 uses to create a database of records that map the global user identifier to the local user identifier.

LIVEINTENT1005, 6:40-46, 8:10-28, fig.2 (reference numeral 36), fig.3 (reference numeral 62). Jaye offers the following example of the special URL:

`http://enterprise_server_id/go?local_server_id&client_information.`

LIVEINTENT1005, 6:40-46.

20. In addition to performing this initial registration process of the local user identifier, the local server collects local session information related to the client 12—including a frequency of page access, a duration of page access, a click stream generated by the user, the subject matter of each web page visited by the user, and other information. LIVEINTENT1005, 7:7-21, 8:29-35, 8:44-10:29, fig.4 (reference numerals 76, 78, 80), fig.5. The local server then uploads the gathered information to the enterprise server 16 with the local user identifier and the local server identifier, and the enterprise server 16 combines this information with information collected from other local servers to create a global user profile associated with the global user identifier. LIVEINTENT1005, 7:21-45, 8:34-52, 10:29-10:41, fig.4 (reference numeral 82), fig.6; *see* LIVEINTENT1005, 10:42-13:18.

21. Jaye describes that “[t]he enterprise server may parse the global user profile and release to an interested party, such as an advertiser, only those portions that are of interest to the advertiser and/or to which the advertiser subscribes.” LIVEINTENT1005, 14:22-25.

B. Robinson-Jaye Combination

1. A POSITA Would Have Been Led to Combine Robinson’s and Jaye’s Teachings.

22. A POSITA would have found it obvious to implement in Robinson a distributed identification and user tracking scheme as taught in Jaye to enable

Robinson's host machines (that host sites and pages of those sites) to: (1) create local user identifiers for locally tracking information about a user; (2) communicate these local user identifiers to Robinson's central system to allow the central system to map the local user identifiers to a single global user identifier or cookie for the user; and (3) store and track information about the user to build a local user profile and upload the local user profile to the central system for building a global user profile for the user.

23. A POSITA would have recognized that Robinson's system would benefit from improvements in view of Jaye's teachings. Robinson teaches Smart Ad Box code that: (1) includes tracking code allowing the central system to access a single cookie stored on the user machine to identify the user, track the user's activities as the user visits tracking-enabled pages hosted by different host machines, and store this tracked information in the central database (LIVEINTENT1004, 2:49-57, 8:40-50, 8:61-9:14, 9:21-37, 10:8-14), and (2) selects and displays advertising targeted to the user based on the tracked information (LIVEINTENT1004, 2:13-67, 4:8-13, 4:24-29, 4:39-67).

24. As a POSITA would have discerned, Robinson describes the central system tracking, storing, and using information about users visiting web pages to generate a global user profile for enabling targeted advertising to be displayed on the pages. But Robinson does not describe mechanisms for a host machine to use

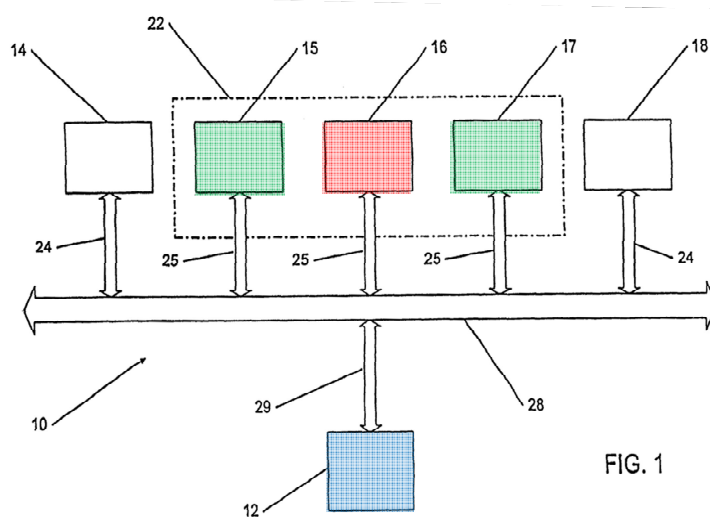
its own identification scheme for reidentifying users that return to pages on the same host machine over multiple sessions. Nor does Robinson describe that the host machine, using such an identification scheme, generates a local user profile from locally tracked information to enable that host machine to tailor non-advertising-related web page content to the user.

25. To remedy these deficiencies, a POSITA would have found Jaye's teachings as to its distributed identification scheme that leverages a global identification scheme in combination with local user identification scheme to be readily applicable to Robinson to yield predictable results. As shown in the diagrams of Robinson and Jaye below, like Robinson, Jaye teaches a system that involves a **central enterprise server 16** (analogous to Robinson's **central system**) that collects information to create a global user profile (analogous to the tracked information about a user collected by Robinson's central system) about the **client** (analogous to Robinson's **user machine**) as the client visits pages hosted by disparate **local servers** (analogous to Robinson's **host machines**).

LIVEINTENT1005, 3:49-50, 3:54-55, 8:40-52, 9:5-41, 10:22-41. Jaye and Robinson teach tracking similar information about users as the users visit web pages, including duration of time spent on a page, user clicks, frequency of page visits, nature of content on web pages, and a user's interest in particular categories. LIVEINTENT1004, 2:31-38, 9:21-35, 15:16-18:59; LIVEINTENT1005, 7:7-26,

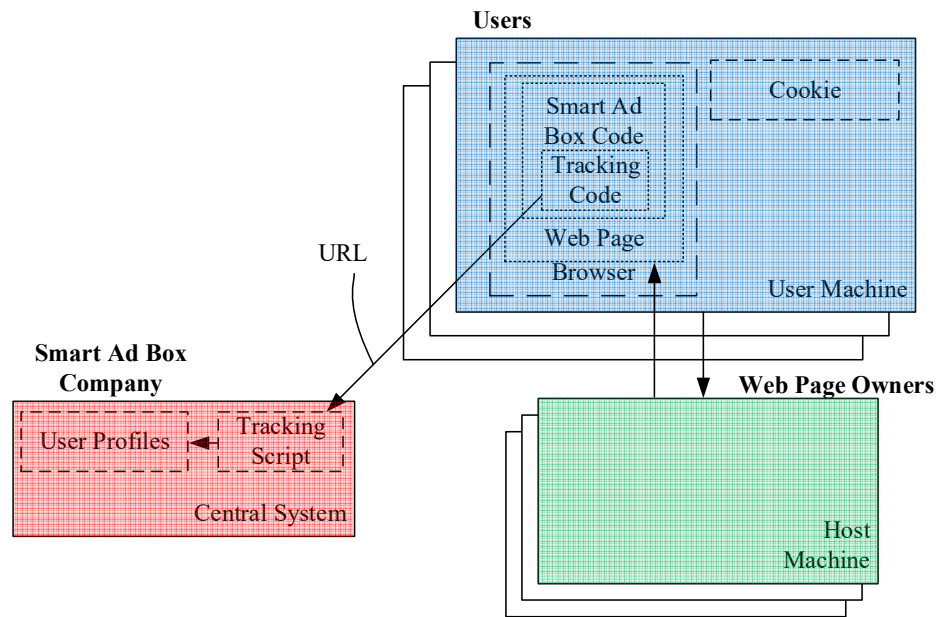
8:29-52, 9:27-10:11. Furthermore, Robinson and Jaye both disclose the central system or server using a cookie to identify the user for creating a user profile about that user. LIVEINTENT1004, 9:12-16, 9:21-35, 9:36-53, 10:8-14; LIVEINTENT1005, 6:40-46, 8:10-28. In view of at least these similarities, a POSITA would have recognized Jaye's teachings to be readily applicable to Robinson to yield predictable results (i.e., a system that offers Robinson's targeted advertising based on globally tracked user activity while also offering Jaye's locally tracked user activity and Jaye's locally provided web page content tailored based on this locally tracked user activity).

Petitioner-Provided Diagram 2: Annotated Figure 1 of
Jaye¹



¹ Throughout this Declaration, annotations to figures and diagrams are shown in color.

Petitioner-Provided Diagram 3: Robinson's System



26. In more detail, a POSITA would have been led to implement Jaye's distributed user identification and tracking scheme in Robinson to enable local host machines to track and store a user's activities in local user profiles associated with local user identifiers, thereby beneficially allowing a given host machine to reidentify users that return to pages hosted by the host machine to create the local user profiles and to use the local user profiles to tailor non-advertising web page content to that user. LIVEINTENT1005, 1:28-45, 1:62-2:11. This distributed user identification and tracking scheme would additionally have maintained the central system's ability to track a user's global activity, i.e., the user's activity across multiple host machines, for the purpose of serving more effective targeted advertisements, as both Robinson and Jaye teach. LIVEINTENT1004, 2:28-31,

2:49-51, 4:9-14, 4:61-6:5; LIVEINTENT1005, 2:1-14, 3:49-50, 3:54-55, 8:40-52, 9:5-41, 10:22-41, 14:22-25. As a result, in Robinson-Jaye, both the host machines and the central system could use the information in the local user profiles (including information consistent with the information tracked about users taught in both Robinson and Jaye) to deliver targeted content (e.g., including both non-advertising and advertising content) to users. *See* LIVEINTENT1004, 2:31-38, 9:21-35, 13:65-14:8, 15:36-50, 16:53-61, 17:2-4; LIVEINTENT1005, 2:6-10, 7:7-36, 8:29-52, 9:27-10:11, 14:22-25.

27. Combining Robinson's and Jaye's teachings also would have involved combining known prior art elements (Robinson's Smart Ad Box system for tracking and storing global user activity to present targeted advertising and Jaye's system for tracking and storing local user activity for the purpose of tailoring web page content) according to known methods (storing and tracking local user identifiers in the central system via special URLs, using the local user identifiers to track local user activity to build local user profiles, and uploading the local user profiles to the central system to generate a global user profile as Jaye teaches) to yield predictable results (the host machines storing and tracking local user identifiers, providing the central system with the local user identifiers via the special URL, and sending the local user profiles to the central system to generate a global user profile for advertisement targeting).

28. A POSITA would have had a reasonable expectation of success combining Robinson and Jaye at least because Robinson and Jaye disclose similar network architectures, web page configurations, and methods of tracking users.

29. First, as discussed above, Robinson's and Jaye's networks both include a central server (e.g., Robinson's central system and Jaye's enterprise server) connected to multiple local servers (e.g., Robinson's host machines and Jaye's local servers). LIVEINTENT1004, 8:61-9:20; LIVEINTENT1005, 4:5-31.

30. Second, Robinson and Jaye both disclose tracking user information as a user interacts with a page on a local server or host machine. LIVEINTENT1004, 8:61-9:53; LIVEINTENT1005, 4:64-5:19, 6:12-7:6, 9:5-10:41.

31. Third, Robinson and Jaye both disclose using a global user identifier (e.g., a cookie) to allow a central server or system to identify a user for creating a global profile of information about the user. LIVEINTENT1004, 9:12-16, 9:21-35, 9:36-53, 10:8-14; LIVEINTENT1005, 6:40-46, 8:10-28.

32. Fourth, Robinson and Jaye both disclose using URLs on pages hosted by host machines to point a user's browser to the central system. As Robinson describes, each time a user references a tracking-enabled page, the Smart Ad Box code on the page includes a Tracking script "which exists on the machine containing the central database" and "calls on the same Tracking Script ... accessible through the one URL." LIVEINTENT1004, 9:12-13, 9:47-53. And Jaye

discloses the special URL that “points to the enterprise server”

LIVEINTENT1005, 6:38-40; *see* LIVEINTENT1005, 6:36-46, 6:47-56, 7:52-53.

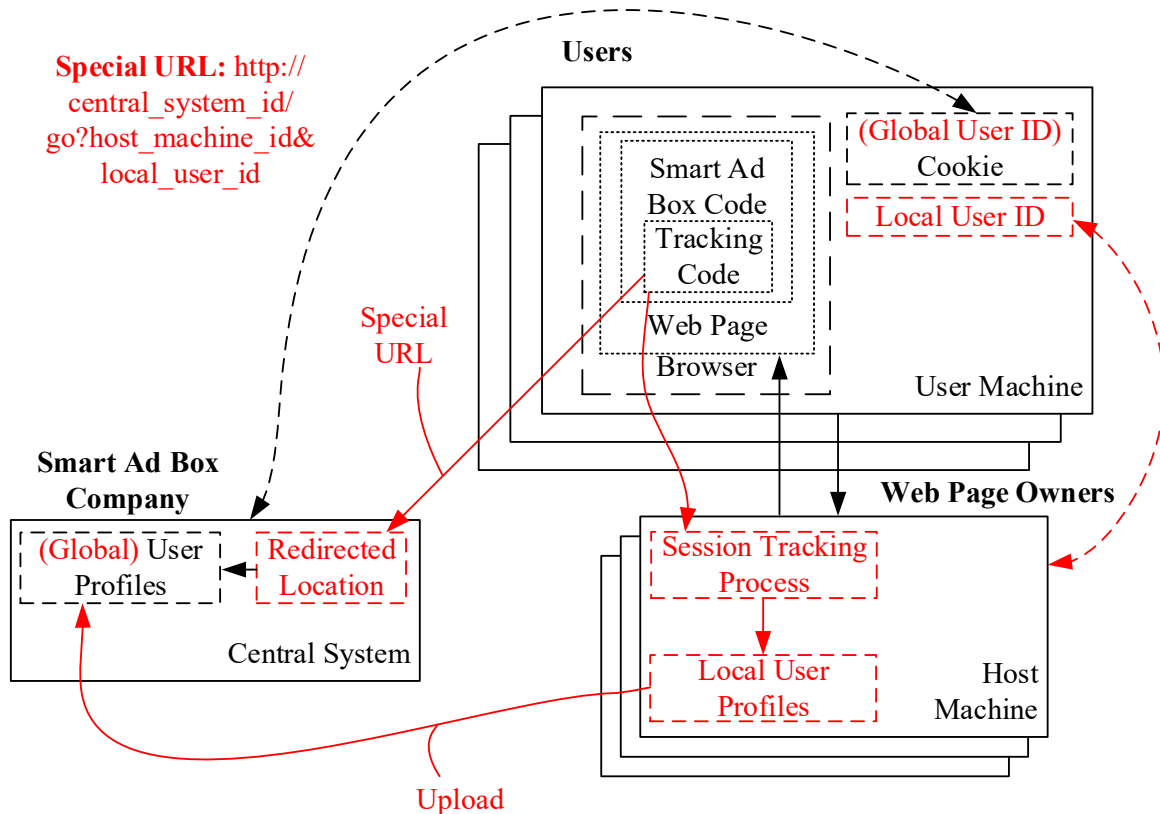
33. Finally, Robinson and Jaye both teach implementing their systems using conventional computer hardware with processes being implemented by the hardware using conventional computer software techniques. LIVEINTENT1004, 8:41-50, 8:61-9:20, 9:49-53, 19:27-33; LIVEINTENT1005, 4:5-23, 13:42-62.

34. For these reasons, a POSITA would have reasonably expected to be successful in combining the teachings of Robinson and Jaye.

2. The Combined Robinson and Jaye System

35. The diagram below depicts the combined Robinson-Jaye system for achieving the benefits articulated above, with modifications to Robinson’s system indicated in red.

**Petitioner-Provided Diagram 4: Robinson-Jaye
Combination**



As shown in this diagram, a POSITA would have combined Robinson's and Jaye's teachings in the following ways.

First, a POSITA would have implemented in Robinson a local user identification scheme as taught in Jaye.

36. Consistent with Jaye's teachings, a POSITA would have configured Robinson's host machine to create a local user identifier on a user machine and configured Robinson's tracking code to use a special URL to force a transfer of the

user machine to the central system to register the local user identifier with the central system. LIVEINTENT1005, 2:6-10, 5:49-6:46, 7:21-26, 8:10-39. The host machine would create the local user identifier only when the user machine first accesses a page hosted by the host machine. LIVEINTENT1005, 5:49-5:67, 6:17-46, 8:29-39. The local user identifier created by the host machine would allow the host machine to reidentify the user machine—and hence the user—in subsequent visits, thereby allowing content to be tailored to the user. LIVEINTENT1005, 5:49-6:31.

37. As Jaye teaches, to register the local user identifier with the central system, the host machine would force the transfer of the user machine to the central system using the special URL implemented in Robinson’s tracking code to enable the central system to map a local user identifier and a host machine identifier (“site identifier for the local server” or “information identifying the local server,” LIVEINTENT1005, 7:31, 8:25-26) to the global user identifier. LIVEINTENT1005, 2:6-10, 5:49-6:11, 6:40-46, 7:21-26, 7:46-7:53, 8:10-39. The special URL would be constructed to enable the central system to access each of these identifiers to generate the mapping by (1) redirecting the user machine to the central system to allow the central system to access the global user identifier on the user machine. (LIVEINTENT1005, 6:32-46, 7:46-8:14) and (2) including as

appended information the local user identifier and the host machine identifier (LIVEINTENT1005, 6:32-46, 7:46-52).

38. In accordance with both Robinson's and Jaye's teachings, the cookie identifying the user machine to the central system taught in Robinson would function as a global user identifier that is known to only the central system (thereby addressing the previously noted privacy concerns) and that allows information collected about the user to be correlated with a single identifier. LIVEINTENT1004, 8:40-50; LIVEINTENT1005, 5:51-5:56, 6:6-11. In this regard, as both Robinson and Jaye teach, if the central system cannot identify a global user identifier on the user machine (e.g., because the user machine has not previously accessed the central system or because the global user identifier has expired), the central system would assign the user a global user identifier (or a cookie that would contain information corresponding to the global user identifier). LIVEINTENT1004, 10:8-29; LIVEINTENT1005, 7:58-8:28.

Second, a POSITA would have implemented in Robinson a process for creating local user profiles as taught in Jaye.

39. Consistent with Jaye's teachings, a POSITA would have configured Robinson's host machine to, after the local user identifier and the host machine identifier are correlated with the global user identifier, create a local user profile associated with the local user identifier and upload this local user profile with the

local user identifier and the host machine identifier to the central system for storing in the user's global user profile. LIVEINTENT1005, 3:49-50, 3:54-55, 7:7-31, 8:40-52, 9:5-41, 10:22-41; *see also* LIVEINTENT1004, 9:17-20 (Robinson contemplating a process in which the host machine includes the tracking script to collect the tracked information and then communicates the tracked information to the central system).

40. As Jaye teaches, the host machine would initiate a local session tracking process to create the local user profile of the user by tracking user activity in a user's local session in which the user visits pages hosted by the host machine. LIVEINTENT1005, 3:49-50, 7:7-31, 8:40-52, 9:5-41, fig.2 (showing that at step 42, a local user profile is established), fig.4 (showing at steps 72, 74, 76, 78, 80, local profile information is collected and stored in a local database). The host machine would subsequently upload the local user profile to the central system so that the central system could update a global user profile identified by the user's global user identifier. LIVEINTENT1005, 3:54-55, 8:40-52, 10:22-41, fig.4 (showing at step 82, the local profile information is uploaded to the enterprise server). The host machine would upload the local user profile with the local user identifier and the host machine identifier, thereby enabling the central system to identify the global user profile associated with the local user profile by using the previously created mapping between the local user identifier and the global user

identifier. LIVEINTENT1005, 6:6-11, 7:21-26, 8:40-52; *see* LIVEINTENT1005, 2:6-10, 3:6-9, 3:17-21, 3:51-55, 5:5-19, 5:49-6:11, 7:19-36, 8:10-39, fig.2 (showing that step 42 in which local user profile is established occurs after step 36 in which the user is forcibly transferred to enterprise server), fig.3 (showing that the local user identifier is mapped to the global user identifier at step 62, which occurs as part of step 36 in fig.2), fig.4 (showing that the local user profile is uploaded to the enterprise server at step 82, which occurs as part of step 42 in Figure 2). As depicted in the diagram of Robinson-Jaye above, a POSITA would have recognized that the user profile maintained by Robinson's central system functions as a global user profile, as taught in Jaye. LIVEINTENT1005, 3:54-55, 8:40-52, 10:22-41.

41. The information tracked by the host machine and then uploaded to the central system would include information consistent with the tracked information taught in both Robinson and Jaye, including: (1) information about the site or page visited by the user, (2) information about an advertisement presented to the user in the Smart Ad Box, (3) information about a purchase made by the user on site or page, and (4) information about other characteristics of the user. *See* LIVEINTENT1004, 2:31-48, 3:19-23, 3:35-39, 4:30-37, 6:53-7:25, 8:16-20, 9:21-35, 13:65-14:8, 14:55-58, 15:29-42, 16:3-8, 16:53-58, 17:2-3; LIVEINTENT1005, 7:7-36, 8:29-52, 8:62-65, 9:27-10:11, 10:36-39, 13:19-31. For example:

- a. Information about the site or page visited by the user would include:
 - (1) a site or host machine identifier of the site visited by the user,
 - (2) page identifier(s) of the page(s) visited by the user, (3) the duration of time spent on a page or site, (4) user clicks on the page or site, (5) frequency of page or site visits, (6) the nature, content, and interest categories of the visited web sites and pages, and (7) any ratings supplied by the user on the site or page. LIVEINTENT1004, 2:31-48, 9:21-35, 13:65-14:8, 15:36-42; LIVEINTENT1005, 7:7-36, 8:29-52, 9:27-10:11.
- b. Information about an advertisement presented to the user in the Smart Ad Box would include: (1) the identifier of the advertisement displayed in the Smart Ad Box, (2) whether the user clicked on the advertisement, (3) whether the user rejected the advertisement or expressed disinterest in an advertisement, and (4) a number of a user's impressions of an advertisement before the user clicked on the advertisement. LIVEINTENT1004, 2:44-48, 9:29-33, 17:2-3; *see* LIVEINTENT1004, 5:4-9.
- c. Information about a purchase made by the user would include: (1) an identifier of any purchased items, (2) the price of items on the host machine's page or site, and (3) an amount of money spent by the user

while visiting the page or site. LIVEINTENT1004, 2:39-41, 9:25-27, 9:34-35.

- d. Information about other characteristics of the user would include demographic and geographic information about the user.

LIVEINTENT1004, abstract, 3:19-23, 3:35-39, 4:30-37, 6:53-7:16, 7:18-25, 8:16-20, 14:55-58, 15:29-35, 16:3-8; LIVEINTENT1005, 8:62-65, 10:36-39, 13:19-31.

42. In implementing Robinson-Jaye, a POSITA would have included in the uploaded profile information the host machine identifiers, as both Robinson and Jaye teach. LIVEINTENT1004, 2:31-38, 13:65-14:8, 15:36-42; LIVEINTENT1005, 7:27-36. A POSITA would have done this for several reasons. Owners of host machines could use host machine identifiers to determine their compensation for displaying advertisements using the Smart Ad Box on pages of their sites (e.g., as captured in the information about the advertisement in the uploaded information). *See* LIVEINTENT1004, 13:58-14:8. The central system could track an accumulation of compensation to which a host machine's owner is entitled for display of the advertisements. *See* LIVEINTENT1004, 14:33-41. The central system could track the nature of content for both the site and the pages within the site (including tracking the subject matter of the site and the pages and the typical characteristics of users who visit the site and the pages) and further

determine the interests and other characteristics of users based on this information. *See* LIVEINTENT1004, 2:35-48, 15:30-50; LIVEINTENT1005, 8:40-9:4, 9:42-10:11, 10:22-41, 11:45-13:18, figs.5-6. The host machine identifier is, for example, a site identifier or a URL of a site, and, therefore, the host machine identifier identifies both the host machine and the site, making inclusion of the host machine identifier not only useful in identifying the host machine, but also in identifying the site and its associated content. *See* LIVEINTENT1004, 2:31-38, 13:65-14:8, 15:36-42; LIVEINTENT1005, 4:53-63, 8:23-26.

43. In implementing Robinson-Jaye, a POSITA would have stored the local profile information in any conventional manner. *See* LIVEINTENT1005, 8:16-28; LIVEINTENT1007, 7:34-54, 9:22-26. A POSITA would have found it obvious to implement Robinson-Jaye such that the stored local profile information is organized using data structures where different information in the stored local profile information are stored in related but distinct records, e.g., using arrays, indices, tables, or other appropriate data structure, to achieve improved storage efficiency and speed. *See* LIVEINTENT1005, 8:16-28; LIVEINTENT1007, 7:34-54, 9:22-26.

44. Moreover, a POSITA would have implemented Robinson-Jaye such that the central system determines whether the local profile information contains new interest categories of a site or page visited by the user that were not previously

contained in the user's global user profile as part of its process for determining a user's interest profile. Specifically, in Robinson-Jaye, consistent with Jaye's teachings, the uploaded local profile information includes information about the site or page visited by the user, including the nature, content, and interest categories of the visited site or page. *See supra* ¶¶41-42. To identify user interests, the central system without human intervention processes this information to determine a user's interest profile in the global user profile, determines "if there has been a previously determined interest score, *i*, for the selected interest category," and computes an initial interest score for a category "[i]f no such historical score exists for the interest category of the user." LIVEINTENT1005, 10:42-11:51, 12:17-23, fig.6; *see* LIVEINTENT1005, 12:17-43. The process executed by the central system, as Jaye teaches, "generates a user profile and/or updates an existing user profile for a specific interest category in a single pass," and the "user profile can be incremented in real time and new interest categories can be easily added." LIVEINTENT1005, 13:32-36.

45. The central system in Robinson-Jaye would, in accordance with Jaye's teachings, "take into consideration the age of the collected behavior information, so that older behavior information has less impact on an interest score than more recent behavior information." LIVEINTENT1005, 8:55-59; *see* LIVEINTENT1005, 2:58-62. The central system's process for determining an

impact of the interest category of the page visited by the user on the calculated interest score for the user involves scaling a value derived from a content interest score of the interest category of the visited page (variable c in Equation 5 of Table 2) by an exponential function ($e^{-\left(\frac{\ln(2)}{\gamma}(t_c-t)\right)}$ in Equation 5 of Table 2) that accounts for an age (e.g., a difference between a current time and a time that the user accessed the page) of the information collected about the user's visit to the page. LIVEINTENT1005, 12:38-41; *see* LIVEINTENT1005, 9:45-10:6, 11:45-12:16, 12:23-42, tbl.1, tbl.2 (equation 5). A POSITA would have implemented Robinson-Jaye to indicate, as Jaye teaches, an age associated with the behavior information for the user in the uploaded local profile information to improve the central system's ability to accurately determine the user's interests and to serve effective targeted advertisements. LIVEINTENT1005, 2:58-62, 8:55-59, 12:16-23, 12:59-65; *see supra* ¶¶22-34; *infra* ¶¶46-50. A POSITA would have understood and/or found obvious that the age associated with the behavior information corresponds to a credibility of the behavior information, with older information being less credible and hence having less of an impact on computed interests for users. *See* LIVEINTENT1005, 2:58-62, 8:55-59, 12:16-23, 12:59-65.

Finally, a POSITA would have implemented Robinson-Jaye to include a process for presenting targeted advertising to users.

46. A POSITA would have implemented Robinson-Jaye such that, when the user visits a page, the Smart Ad Box would appear on the page and the user would see an advertisement targeted to the particular user. LIVEINTENT1005, 4:43-46. The central system would use the uploaded local profile information collected from the host machines, as Robinson teaches, to determine the advertisement to present to the user in the Smart Ad Box. LIVEINTENT1004, 2:23-51, 3:64-4:6, 4:51-64, 15:14-62, 16:49-18:67. Indeed, Jaye confirms that the central system can “parse the global user profile and release” portions of the information to an advertiser. LIVEINTENT1005, 14:22-25.

47. Specifically, a POSITA would have implemented in Robinson-Jaye the user tracking process in Jaye—including the process for identifying user interests based on the interests of the pages visited by the user and adding these user interests to the global user profile—to provide more effective targeted advertising to the user. LIVEINTENT1005, 10:42-46, 11:45-51; *see* LIVEINTENT1005, 2:50-62, 9:42-12:65, figs.5-6; *supra* ¶¶22-34. The central system would use the uploaded local profile information to determine user interests, as taught in both Robinson and Jaye. LIVEINTENT1004, 15:16-28, LIVEINTENT1005, 10:42-46, 11:45-51; *see* LIVEINTENT1004, 1:27-44, 2:9-22, 15:16-16:67; LIVEINTENT1005, 9:42-12:65, figs.5-6. And the central system would further use the determined user interests to identify targeted advertising to

present to the user, as Robinson teaches. LIVEINTENT1004, 15:16-28; *see* LIVEINTENT1004, 1:27-44, 2:9-22, 2:23-51, 3:64-4:6, 4:51-64, 15:14-18:67; *see also* LIVEINTENT1005, 14:22-25 (Jaye indicating that advertisers would have an interest in information in a global user profile).

48. A POSITA would have implemented Robinson-Jaye so that the central system uses the page identifier to: (1) determine the subject matter of the page and use the subject matter of the page for selecting an advertisement, and (2) identify whether an advertiser permits an advertisement to be presented on the page visited by the user. Robinson teaches tracking the page identifier of the page visited by the user, and Jaye teaches using page identifiers to track information about the content of pages. LIVEINTENT1004, 9:21-23; LIVEINTENT1005, 9:42-10:41, fig.5. A POSITA would have been motivated to use the page identifier of the page including the Smart Ad Box that presents the advertisement to the user in order to identify the page, determine the subject matter of that page, and then select an advertisement that would have “resonance” with and “relate in some way to the subject matter of the page,” thereby improving the central system’s ability to select an advertisement that the user is more likely to interact with.

LIVEINTENT1004, 4:64-67. Alternatively/additionally, as Robinson teaches, advertisers can select where their advertisements are displayed by indicating to the central system which sites and pages a particular advertisement should be

displayed, and a POSITA would have implemented Robinson-Jaye so that the central system uses the page identifier of the page visited by the user to determine whether an advertiser has identified the visited page as a desirable page or undesirable page for displaying an advertisement. LIVEINTENT1004, 3:29-34, 5:11-21, 5:30-6:15.

49. As to advertisement identifiers, a POSITA would have implemented Robinson-Jaye such that the advertisement identifier is used for selecting and presenting an advertisement to the user. Robinson teaches that the advertising identifier is used for selecting an advertisement to present to the user and teaches tracking the identifier of the advertisement that was displayed to the user in the Smart Ad Box. LIVEINTENT1004, 5:37-40, 5:59-63, 6:4-12; *see* LIVEINTENT1004, 5:37-6:15 (describing the advertisement identifier is used for selecting an advertisement to present on a Smart Ad Box), 9:8-35 (describing that the identifier of the advertisement displayed to the user in the Smart Ad Box is tracked as the user visits a page).

50. The Robinson-Jaye combination would therefore enable host machines to build and use local user profiles to tailor non-Smart Ad Box content to the users while also enabling the central system to build global user profiles capturing the interests of users to thereby provide more effective targeted

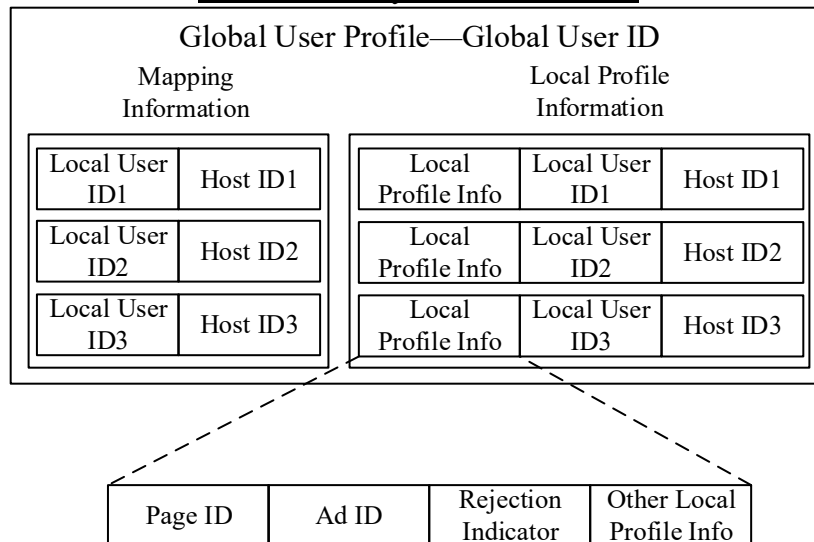
advertisement through the Smart Ad Box. LIVEINTENT1004, 4:9-14, 9:21-35, 15:16-16:67; LIVEINTENT1005, 2:3-10, 9:27-10:21, 10:42-13:41.

3. In the Robinson-Jaye Combination, a Global User Profile Stored in the Central System is Constructed from Local User Profiles Stored on Multiple Different Host Machines.

51. The diagram below shows an example of a global user profile collected by the central system consistent with Robinson-Jaye. As discussed above, the central system stores mapping information that correlates local user identifiers and host identifiers to the specific global user identifier associated with the global user profile by using the forced transfer process mediated by the special URL, as Jaye teaches. LIVEINTENT1005, 7:46-49, 8:10-14, fig.3 (reference numeral 62); *see* LIVEINTENT1005, 2:6-10, 2:21-24, 3:6-9, 3:17-21, 3:34-38, 3:51-53, 5:49-6:11, 6:40-46, 7:21-26, 7:46-7:53, 8:10-39. Moreover, the central system stores local profile information uploaded by the host machines to the central system, which as discussed above, would be uploaded to and compiled by the central system with the local user identifier and the host machine identifier. LIVEINTENT1005, 3:54-55, 7:27-45, 8:29-52, 10:22-41. As Jaye teaches, this global user profile is associated with a single global user identifier (Global User ID) stored on a single user machine and, as depicted in the example diagram below, would represent visits to different host machines (Host Machines 1, 2, and 3), each having a different host machine identifier (Host ID1, Host ID2, Host ID3)

and a distinct local user identifier (Local User ID1, Local User ID2, Local User ID3) for the user machine. LIVEINTENT1005, 2:6-10, 5:49-6:11, 6:40-46, 7:21-26, 7:46-7:53, 8:10-39. As indicated in the diagram below, each set of local profile information provided by a corresponding host machine includes numerous pieces of information, including a page identifier, an advertisement identifier, a rejection indicator, and other local profile information. *See supra* ¶¶40-45.

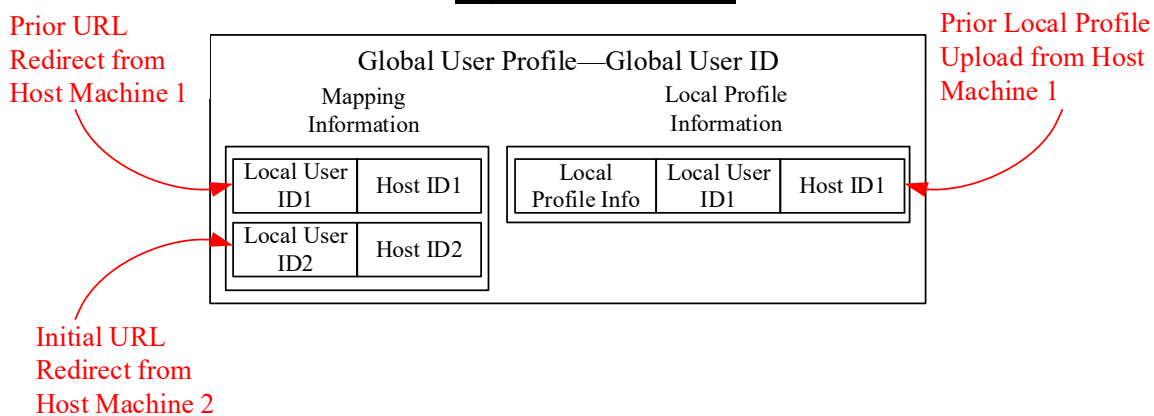
Petitioner-Provided Diagram 5: Global User Profile in Robinson-Jaye Combination



52. By way of example, as depicted in the diagram below, when the user for this global user profile first visits a page hosted by Host Machine 2, the central system would receive Local User ID2 and Host ID2 via the special URL and would then store Local User ID2 and Host ID2 to correlate these identifiers with Global User ID (e.g., using the information from the special URL). *See supra* ¶¶36-38. In this example, Global User ID is assumed to have been previously

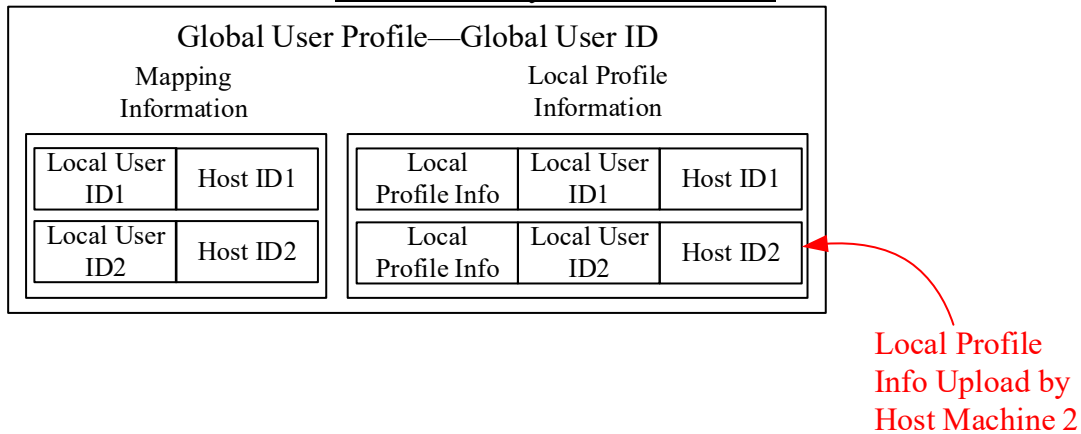
generated (e.g., on the user's prior visit to Host Machine 1 corresponding to Host ID1), and the central system is assumed to have previously collected Local Profile Information (e.g., when Host Machine 1 previously uploaded local profile information). *See supra* ¶¶36-45.

Petitioner-Provided Diagram 6: Generating Mapping Information in the Global User Profile in Robinson-Jaye Combination



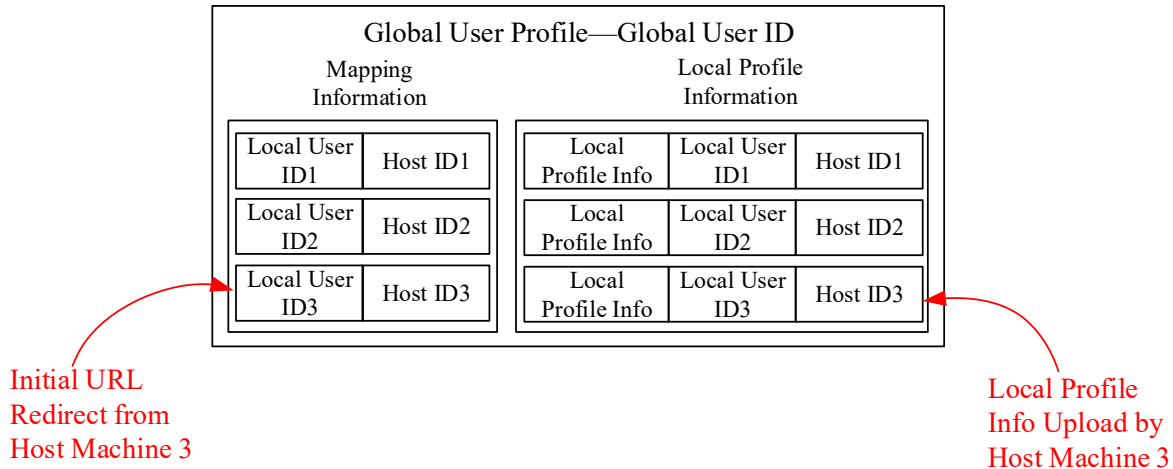
53. Then, after Local User ID2 and Host ID2 are correlated with Global User ID, Host Machine 2 would, consistent with Jaye's teachings, upload local profile information collected over one or more sessions and/or visits by the user to Host Machine 2 and would forward this information with Local User ID2 and Host ID2. LIVEINTENT1005, 3:54-55, 7:27-31, 8:40-52, 10:22-41; *see supra* ¶¶39-45.

**Petitioner-Provided Diagram 7: Collecting Local
Profile Information in the Global User Profile in
Robinson-Jaye Combination**



54. As shown in the diagram below, other host machines, e.g., Host Machine 3 and beyond, can subsequently communicate with the central system and provide local identifiers (Local User ID3, Host ID3) for the central system to correlate with Global User ID to facilitate later upload of local profile information to the central system by these host machines. *See supra* ¶¶36-45. As Jaye teaches, the local profile information associated with each local user identifier and host identifier can be continually updated as the host machines upload session records from the local database of local profile information. LIVEINTENT1005, 3:54-55, 7:27-31, 8:40-52, 10:22-41. A POSITA would have hence understood that the central system would store several different local user profiles as part of a global user profile, and each global user profile would be associated with a different user machine. *See supra* ¶¶36-45.

**Petitioner-Provided Diagram 8: Collecting Additional
Local Profile Information in the Global User Profile in
Robinson-Jaye Combination**



C. An Example of an Implementation of Robinson-Jaye Combination

55. I provide below a technical description of an example of an implementation of the Robinson-Jaye combination consistent with the discussion above.

1. Network Configuration

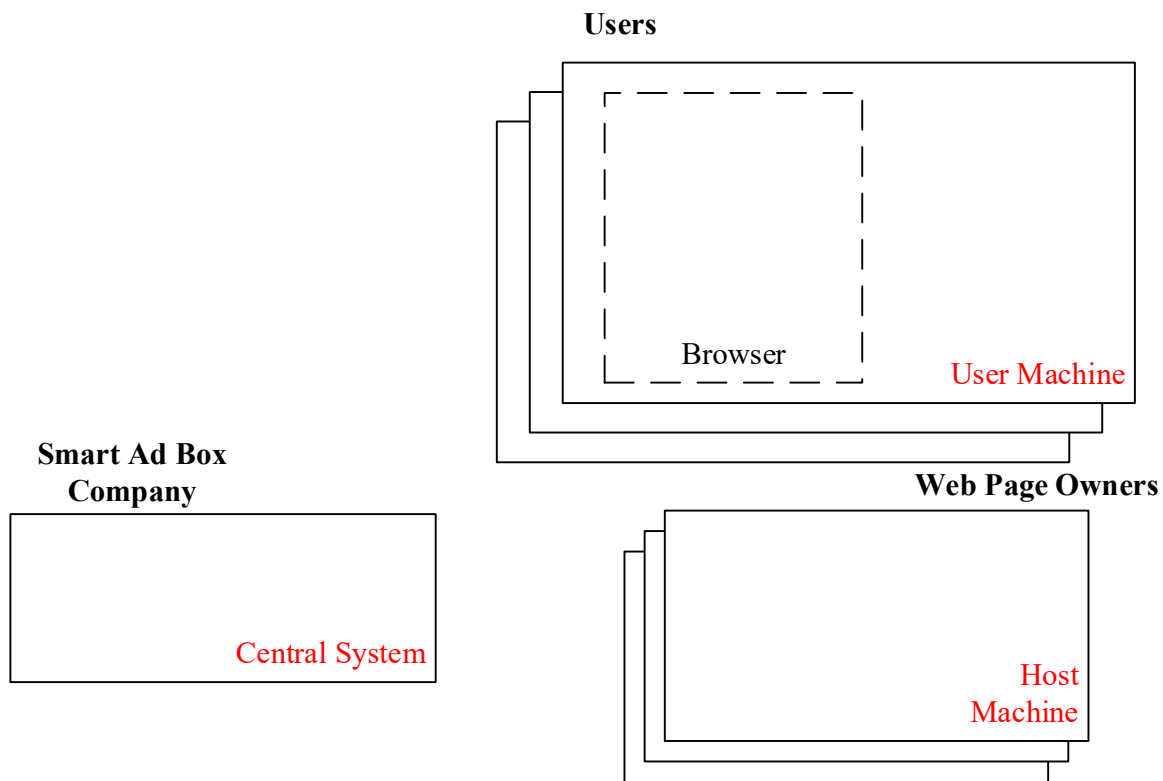
56. The diagram below depicts an exemplary configuration of the network in Robinson-Jaye. The network includes multiple user machines, multiple host machines, and a central system that communicate with one another over the Internet. *See supra* ¶¶25, 29, 35. As discussed above, this example implements a distributed identification and user tracking scheme that simultaneously allows the host machines to generate local user profiles to locally track information about

users and the central system to generate global user profiles to globally track information. *See supra* ¶¶22, 25-26. And the processes performed in the network as described below can be performed by software, application programs used in conjunction with such software, and/or by executing a computer program tangibly embodied in a machine-readable storage device for execution by a processor.

LIVEINTENT1004, 19:27-30; LIVEINTENT1005, 13:42-14:6; *see*

LIVEINTENT1004, 4:8-15.

Petitioner-Provided Diagram 9: Robinson-Jaye System



57. The user machines are operated by users to access tracking-enabled web pages for accessing conventional Internet services and HTML pages using the client. LIVEINTENT1004, 3:61-4:6, 8:41-67, 9:4-45, 10:23-29;

LIVEINTENT1005, 4:7-14, 4:64-66; *see* LIVEINTENT1004, 1:1-3, 1:12-20, 7:52-8:8, 8:21-28; LIVEINTENT1005, 4:45-5:4.

58. The host machines or servers host the tracking-enabled web pages on participating websites. LIVEINTENT1004, 9:49-50; LIVEINTENT1005, 4:46-50; *see* LIVEINTENT1004, 5:48-50, 6:38-42, 9:17-20, 13:22-26, 14:1-8; LIVEINTENT1005, 4:53-5:4, 7:17-26, 8:23-26, 9:42-45, 10:12-21. The host machines (1) are separate from the machine hosting the central database, and (2) host participating web sites that participate in Smart Ad Box services to enable the central database to track the users visiting their pages, that receive payment from the Smart Ad Box company for displaying advertisements using the Smart Ad Box services, and that provide the central database with rules or control records indicating a preference for certain advertisers or advertisements.

LIVEINTENT1004, 4:39-42, 5:10-26, 6:13-37, 8:51-60, 8:64-9:7, 9:12-16, 9:47-53, 13:22-39, 13:58-64; *see* LIVEINTENT1004, 1:59-63 (describing advertisers, ad agencies, and web sites as distinct interested commercial entities);

LIVEINTENT1004, 13:31-34 (“The code [for tracking users] could be designed in such a way that there need be no direct communication between the people supplying the Smart Ad Box service and related services and the people who want to enable their Web site to participate in those services.”). Each host machine implements a local identification scheme that enables the host machine to uniquely

identify users that visit the host machine and to track local profile information accumulated about the users over multiple sessions. LIVEINTENT1005, 2:1-5, 5:20-51, 8:53-55; *supra* ¶¶36-44.

59. The central system includes a central database and a central server machine, which are programmed to be updated with tracked user information as users using the user machines visit tracking-enabled pages. LIVEINTENT1004, 8:40-50, 8:61-9:20, 9:46-10:7; LIVEINTENT1005, 4:5-23; *see* LIVEINTENT1004, 1:12-15, 1:66-2:8, 5:11-13, 8:40-60, 10:35-44, 19:27-33; LIVEINTENT1005, 1:10-14, 1:15-17, 3:10-14, 3:48-49, fig.1. For each user that visits the host machines, the central system populates records of a database that correlate local user identifiers used by the host machines as part of their local identification schemes with a global user identifier and a global user profile associated with the user and tracks the local profile information generated by the host machine in this global user profile. LIVEINTENT1004, 8:40-50; LIVEINTENT1005, 5:51-5:56, 6:6-11, 7:21-26, 8:40-52; *see* LIVEINTENT1005, 2:6-10, 3:6-9, 3:17-21, 3:51-55, 5:5-19, 5:49-6:11, 7:19-36, 8:10-39, figs.2-4; *supra* ¶¶36-40.

2. Registration of Local User Identifier in Global User Profile

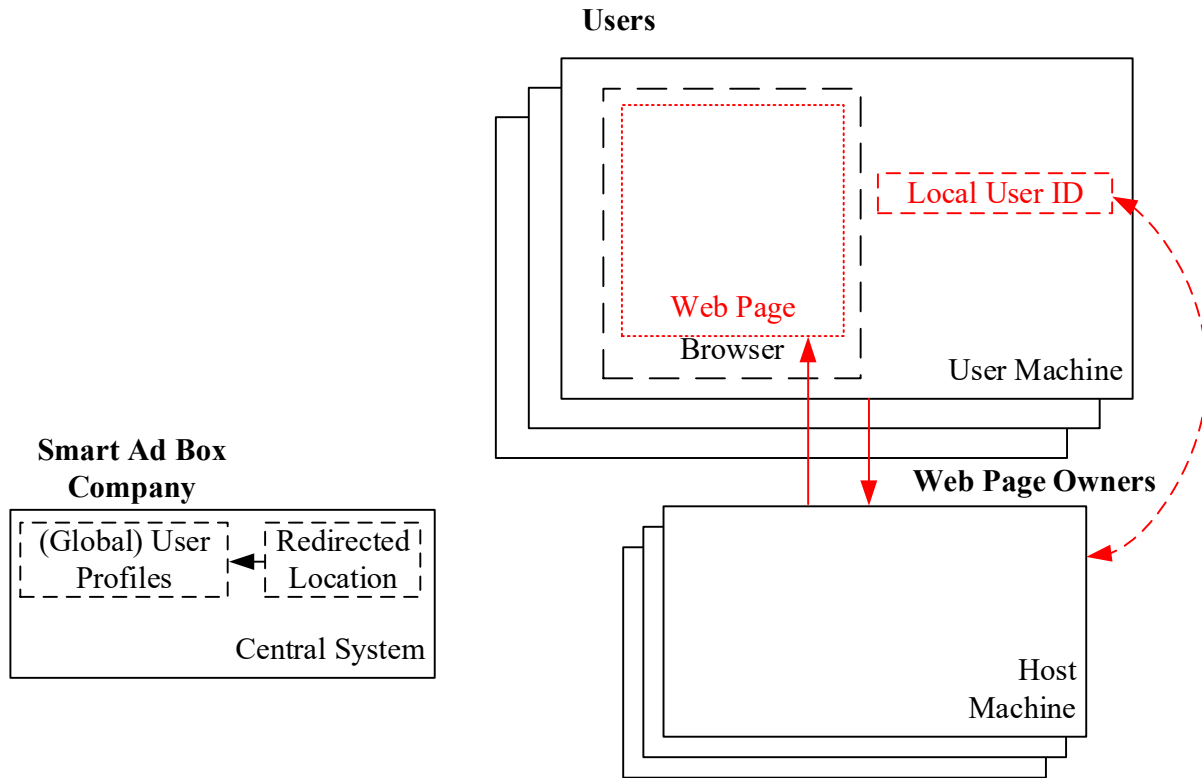
60. The exemplary implementation of Robinson-Jaye includes a process for registering local user identifiers, generated as part of local identification

schemes implemented by host machines, with a global user profile for the user. *See supra* ¶¶36-38. Specifically, the local identification schemes implemented by the host machines involve assigning unique local identifiers to users as the user visit pages of sites hosted by host machines. *See supra* ¶¶36-38.

61. Assume a user using a particular user machine is browsing the Internet and, in particular, accesses one or more tracking-enabled pages of a site hosted by a particular host machine.

62. First, the host machine generates a unique local user identifier for the user. When the user uses a browser on the user machine to access a web page on the site hosted by the host machine for the first time (i.e., the user machine has not previously been used to access the site), the host machine creates the unique local user identifier and stores the local user identifier on the user machine, as shown in the diagram below. LIVEINTENT1005, 5:49-5:67, 6:17-46, 6:67-7:2, 8:29-39. The local user identifier created by the host machine allows the host machine to identify the user machine—and hence the user—in subsequent visits to the web site and hence each time the user machine requests data from the host machine. LIVEINTENT1005, 5:49-6:31. Different host machines would set different local user identifiers for the user machine as the user uses the user machine to visit different sites hosted by these different host machines. LIVEINTENT1005, 5:49-56; *see* LIVEINTENT1005, 5:20-67.

Petitioner-Provided Diagram 10: Robinson-Jaye System

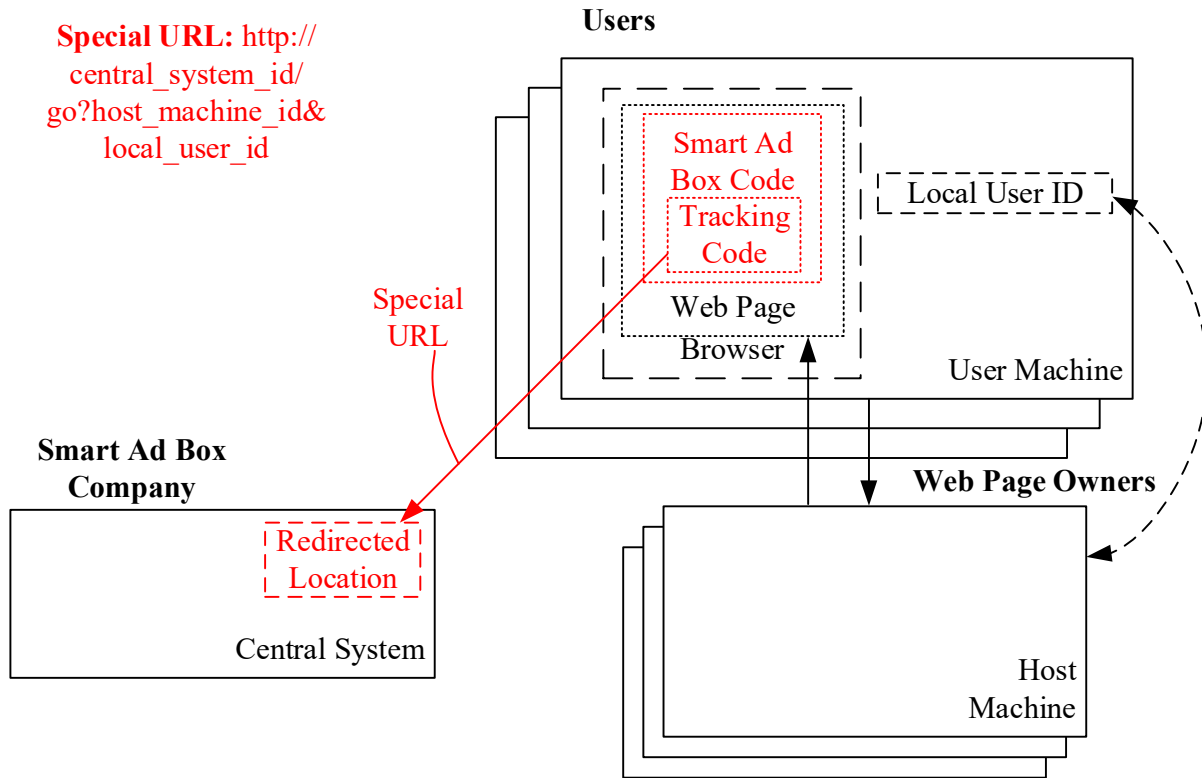


63. Second, the host machine initiates a process for registering the local user identifier assigned to the user with the central system. As depicted in the diagram below, Smart Ad Box code (e.g., HTML code) on the page includes tracking code that allows the central system to use a cookie mechanism to identify the user. LIVEINTENT1004, 2:49-57, 8:40-50, 8:61-9:14, 10:8-14. To enable the central system to use this cookie mechanism, the tracking code contains a special URL that forces a transfer of the user machine from the web page hosted by the host machine to the central system, thereby allowing the central system to identify the user machine and initiate registration of the local user identifier with the central system. LIVEINTENT1005, 2:6-10, 5:49-6:11, 6:40-46, 7:21-26, 7:46-7:53, 8:10-

39; *see* LIVEINTENT1005, 6:32-46, 7:46-8:39, fig.2 (reference numeral 36), fig.3.

The forced transfer serves as a URL redirection that allows the central system to directly receive from the user machine the cookie containing the global user identifier. LIVEINTENT1005, 7:58-63, 8:10-28, fig.3 (elements 54, 62); *see supra* ¶¶36-38. The special URL points the user machine to the central system and thus redirects the user machine to the central system. LIVEINTENT1005, 6:38-6:46; *see* LIVEINTENT1005, 2:43-48, 4:64-5:4, 6:32-46. Additionally, the host machine appends information, including the host machine identifier and the local user identifier, to the end of the special URL. LIVEINTENT1005, 6:40-46, 8:10-28, fig.2 (reference numeral 36), fig.3 (reference numeral 62). As indicated in the diagram below, the special URL is in the form of `http://central_system_id/go?host_machine_id&local_user_id`, though other constructions of the special URL that append other types of information related to the user would have been obvious. LIVEINTENT1005, 6:40-46.

Petitioner-Provided Diagram 11: Robinson-Jaye System

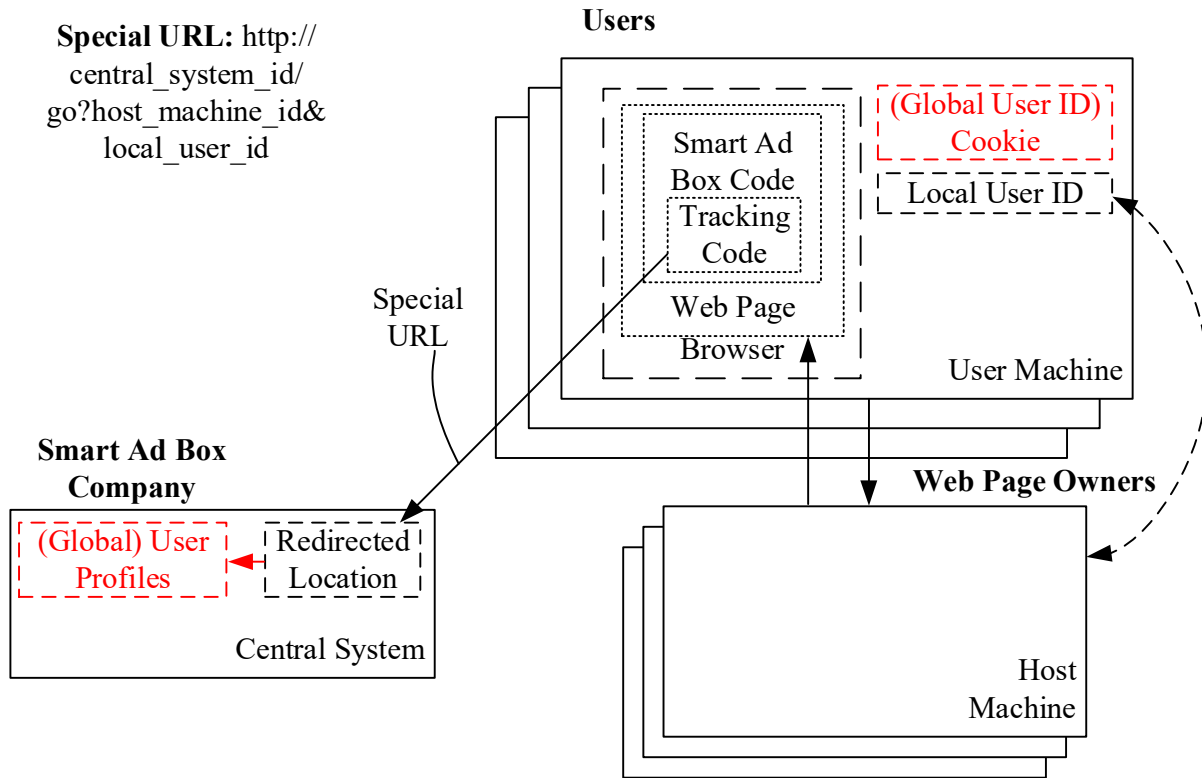


64. The forced transfer gives the central system access to a global user identifier on the user machine. Specifically, as a result of the forced transfer, the central system accesses a global user identifier assigned to the user machine (e.g., a global cookie), as depicted in the diagram below, and hence, as a POSITA would have understood and/or found obvious, receives this cookie without human intervention. LIVEINTENT1004, 9:12-16, 9:21-35, 9:36-53, 10:8-14; LIVEINTENT1005, 6:32-46, 7:46-8:28. If the user machine has not previously accessed the central system, the central system generates the global user identifier, assigns the global user identifier to the user machine, stores the global user identifier on the user machine, and then accesses the global user identifier.

LIVEINTENT1005, 5:49-56, 7:58-8:31; *see* LIVEINTENT1004, 10:18-29;

LIVEINTENT1005, 5:20-67, fig.3 (reference numerals 54, 56, 58, 62). And if the user machine has previously accessed the central system and therefore the global user identifier was previously assigned to the user machine, the central system accesses the global user identifier without generating a new global user identifier for the user. LIVEINTENT1005, 5:49-56, 7:58-8:31; *see* LIVEINTENT1004, 10:8-29; LIVEINTENT1005, 5:20-67, fig.3 (reference numerals 54, 62). The global user identifier functions as a unique global identifier that serves as “persistent state information” assigned to the user machine and can be mapped to multiple local user identifiers set by different host machines for the user machine. LIVEINTENT1005, 5:20-67, 7:58-8:28, fig.3 (reference numerals 54, 62); *see* LIVEINTENT1004, 8:40-50, 9:36-10:29. This global user identifier is associated with a unique global user profile associated with the user where the central system will store information about the user. LIVEINTENT1004, 9:12-16, 9:21-35, 9:36-53, 10:8-14; LIVEINTENT1005, 6:40-46, 8:10-28.

Petitioner-Provided Diagram 12: Robinson-Jaye System

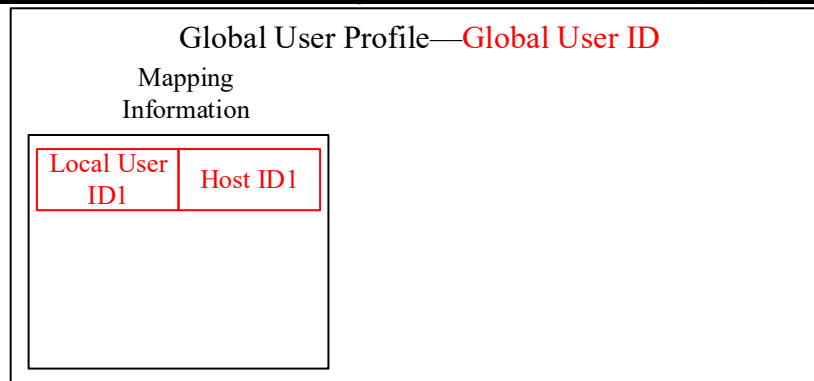


65. Third, the central system then registers the local user identifier and the host machine identifier in the global user profile for the user. For example, as illustrated in the diagram below, the central system registers the local user identifier (Local User ID1) with the central system by mapping the local user identifier and the host machine identifier (Host ID1) (received by the central system in the information appended to the special URL discussed above) to the global user identifier (accessed by the central system as a result of the forced transfer from the host machine to the central system) in the global user profile. LIVEINTENT1005, 2:6-10, 5:49-6:11, 6:32-46, 7:21-26, 7:46-8:39; *see supra* ¶¶63-64. The central system registers the local user identifier and the host machine

identifier in the global user profile associated with the global user identifier in response to the forced transfer and hence without human intervention.

LIVEINTENT1005, 6:32-46, 6:57-58, 7:46-49, 8:10-27, fig.3 (reference numeral 62), fig.4 (reference numeral 82); *see supra* ¶¶36-38.

Petitioner-Provided Diagram 13: Robinson-Jaye System

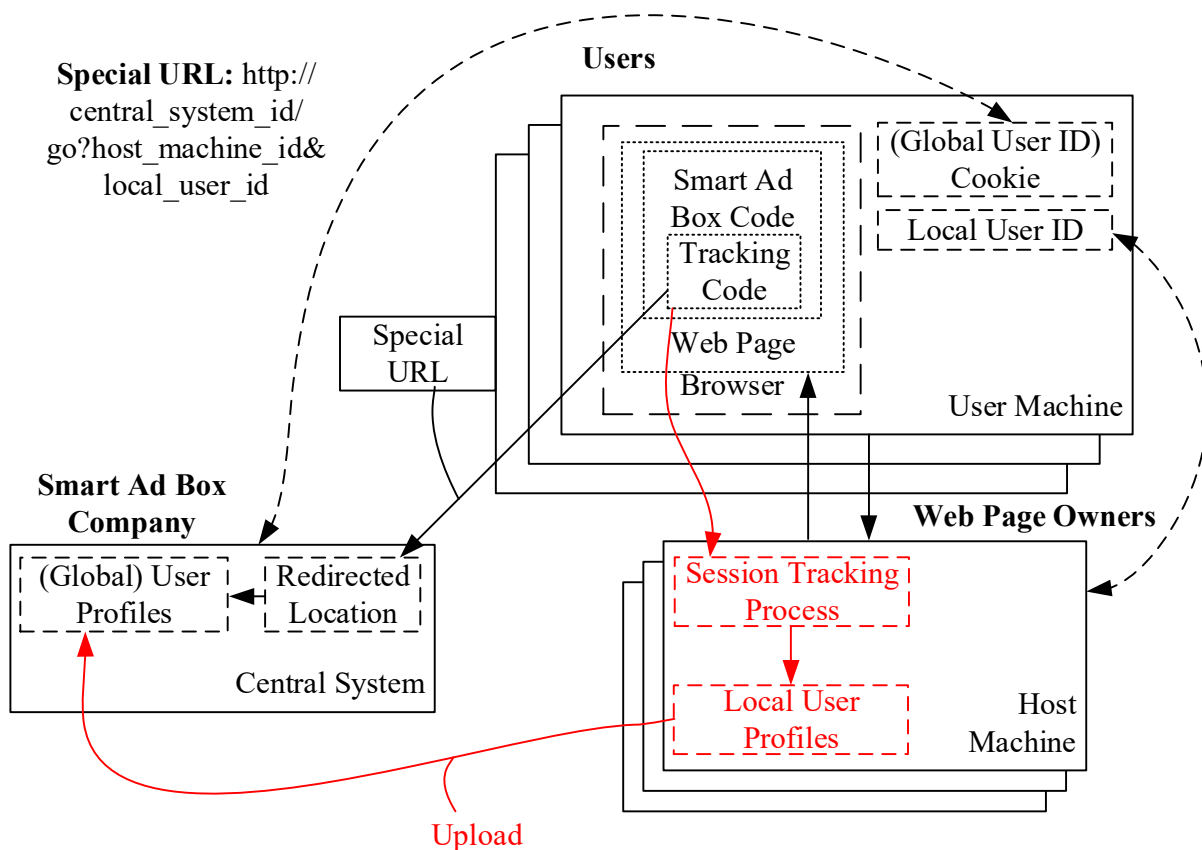


3. Upload of Local Profile Information to Global User Profile

66. Following registration of the local user identifier, the host machine generates local profile information about the user, stores this information in a local user profile associated with the local user identifier, and uploads this information to the central system. *See supra* ¶¶39-42. The host machine initiates a local session tracking process to create the local user profile of the user by tracking user activity in a user's local session in which the user visits pages hosted by the host machine. LIVEINTENT1005, 3:49-50, 7:7-31, 8:40-52, 9:5-41, fig.2 (showing that at step 42, a local user profile is established), fig.4 (showing at steps 72, 74, 76, 78, 80, local profile information is collected and stored in a local database); *see*

LIVEINTENT1004, 9:17-20. The host machine subsequently uploads the local user profile to the central system to update the global user profile (stored on the central system) identified by the user's global user identifier. LIVEINTENT1005, 3:54-55, 8:40-52, 10:22-41, fig.4 (showing at step 82, the local profile information is uploaded to the enterprise server); *see* LIVEINTENT1004, 9:17-20.

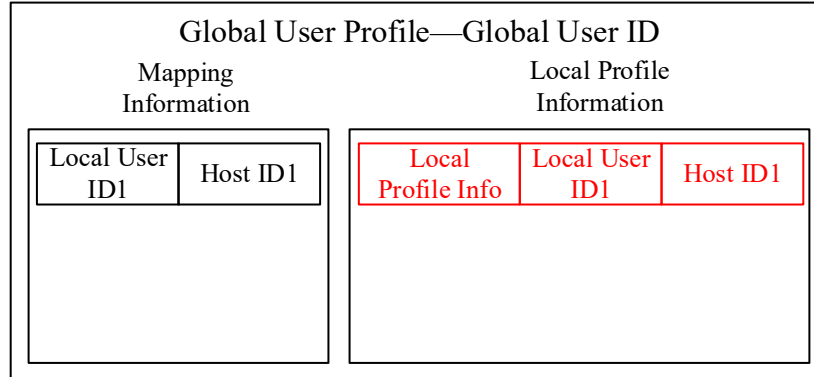
Petitioner-Provided Diagram 14: Robinson-Jaye System



67. As shown in the diagram below, the host machine uploads local profile information (Local Profile Info) in the local user profile with the local user

identifier (Local User ID1) and the host machine identifier (Host ID1), thereby enabling the central system to identify the global user profile associated with the local profile information by using the previously created mapping between the local user identifier and the global user identifier. LIVEINTENT1005, 6:6-11, 7:21-26, 8:40-52; *see* LIVEINTENT1005, 2:6-10, 3:6-9, 3:17-21, 3:51-55, 5:5-19, 5:49-6:11, 7:19-36, 8:10-39, fig.2 (showing that step 42 in which local user profile is established occurs after step 36 in which the user is forcibly transferred to enterprise server), fig.3 (showing that the local user identifier is mapped to the global user identifier at step 62, which occurs as part of step 36 in fig.2), fig.4 (showing that the local user profile is uploaded to the enterprise server at step 82, which occurs as part of step 42 in Figure 2). The central system updates the global user profile with the uploaded local user profile, local user identifier, and host machine identifier in response to the central system identifying the local user identifier and/or the global user identifier (as discussed in greater detail above) and hence without human intervention. LIVEINTENT1005, 3:54-55, 7:27-45, 10:29-41, fig.4 (reference numeral 82); *see* LIVEINTENT1004, 9:36-37, 10:10-14, 10:24-26; *see supra* ¶¶39-42, 51-53.

Petitioner-Provided Diagram 15: Robinson-Jaye System



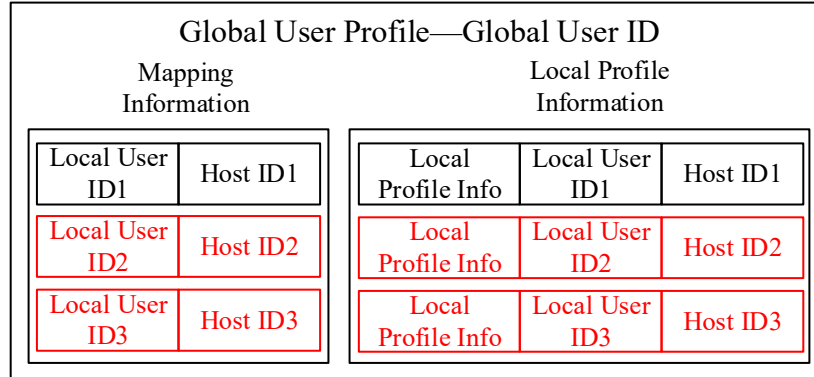
68. Local profile information in the local user profile uploaded to the central system includes the information discussed above in connection with the Robinson-Jaye combination, including (1) information about the site or page visited by the user, (2) information about an advertisement presented to the user in the Smart Ad Box, (3) information about a purchase made by the user on site or page, and (4) information about other characteristics of the user. *See supra* ¶41.

4. Storage of Local User Identifiers and Local Profile Information from Multiple Host Machines and for Multiple Users in Global User Profiles

69. The central system stores mappings of local user identifiers and local profile information from multiple host machines visited by the user and additionally performs a similar process of storing mappings and local profile information for other users that engage with pages and sites hosted by the host machines. For example, as depicted in the diagram below, the global user profile associated with the user's unique global user identifier includes information

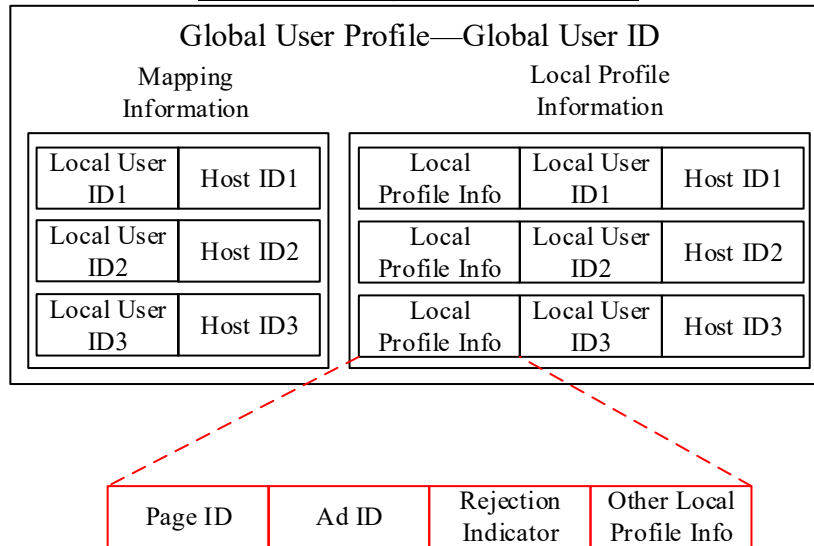
received from additional host machines visited by the user on different occasions. LIVEINTENT1005, 2:6-10, 5:49-6:11, 7:21-26, 8:10-8:28. These additional host machines, like the host machine described above, provide their host machine identifiers (Host ID2, Host ID3) and local user identifiers assigned to the user (Local User ID2, Local User ID3) by using the special URL to initiate the forced transfer to the central system and thereby register the local user identifiers and the host machine identifiers with the global user profile. LIVEINTENT1005, 6:32-46, 7:46-53; *see supra* ¶¶63-65. Additionally, the additional host machines, like the host machine described above, also upload the local profile information—along with the local user identifier and the host machine identifier—that is collected in local user profiles stored on the additional host machines. LIVEINTENT1005, 6:6-11, 7:21-26, 8:40-52; *see* LIVEINTENT1005, 2:6-10, 3:6-9, 3:17-21, 3:51-55, 5:5-19, 5:49-6:11, 7:19-36, 8:10-39, fig.2-4; *see supra* ¶¶66-68. The user can be tracked over multiple sessions, and thus each of the host machines can upload local profile information multiple times. LIVEINTENT1005, 8:53-55; *see* LIVEINTENT1005, 9:22-26, 9:33-38. And the local profile information associated with each local user identifier and host identifier can be continually updated as the host machines upload session records from the local database of local profile information. LIVEINTENT1005, 3:54-55, 7:27-31, 8:40-52, 10:22-41.

Petitioner-Provided Diagram 16: Robinson-Jaye System



70. As indicated in the diagram below, each set of local profile information provided by a corresponding host machine includes numerous pieces of information, including a page identifier, an advertisement identifier, a rejection indicator, and other local profile information. *See supra* ¶¶40-45. The stored local profile information is organized using data structures where different information in the stored local profile information are stored in related but distinct records, e.g., using arrays, indices, tables, or other appropriate data structure. *See supra* ¶43.

**Petitioner-Provided Diagram 17: Global User Profile in
Robinson-Jaye Combination**



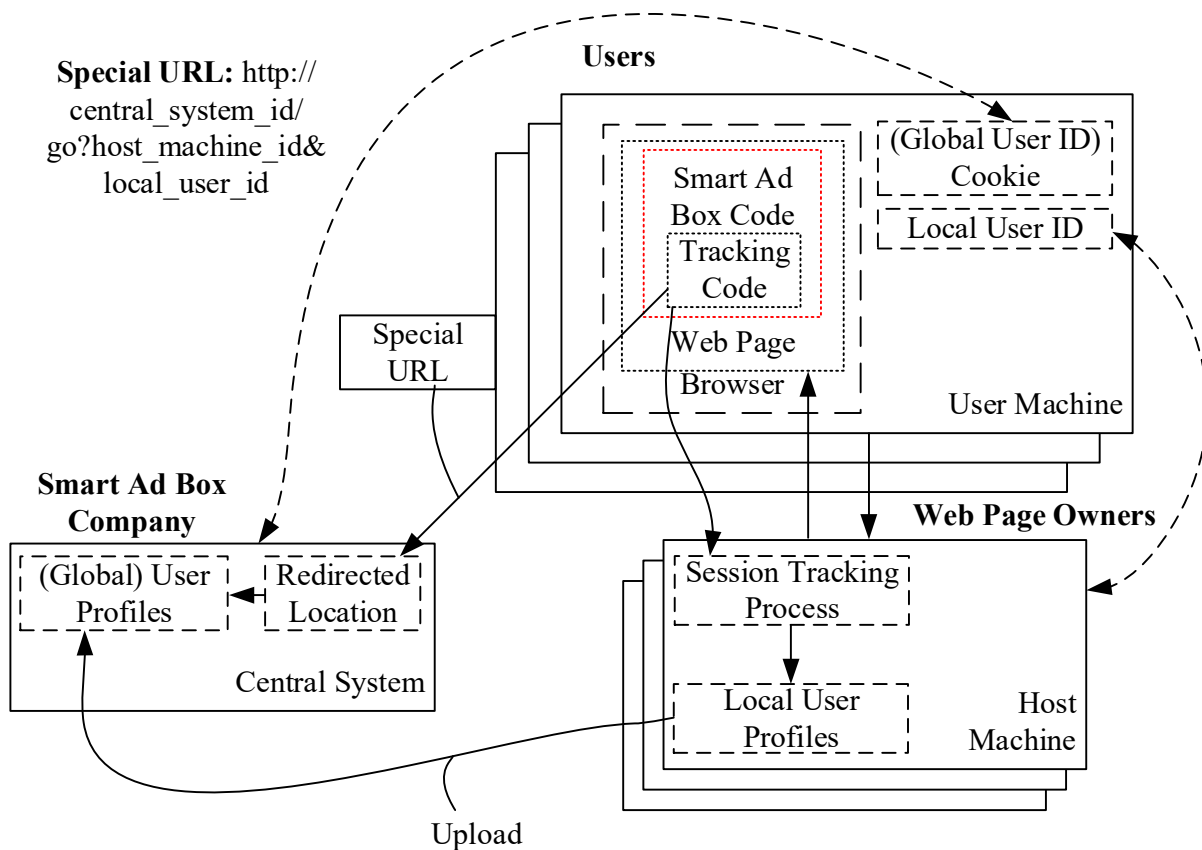
71. The central system includes global user profiles for multiple different users such that the central system generates a unique global user identifier associated with a unique global user profile for each user, registers local user identifiers for each user, and stores local profile information collected about each user, as discussed above. LIVEINTENT1005, 6:6-11, 7:27-31, 11:45-48; *see* LIVEINTENT1005, 8:62-65, 12:61-65; *supra* ¶¶60-69.

5. Targeting Advertising for Users

72. As shown in the diagram below, the Smart Ad Box code selects and displays advertising targeted to a particular user based on tracked information about the user and tracked information about other users. LIVEINTENT1004, 2:13-67, 4:8-13, 4:24-29, 4:39-67. The Robinson-Jaye system provides processes for generating targeted advertising for a user that involve using (1) the local profile

information in the user's global user profile, (2) the local user identifiers in the user's global user profile, and (3) specific pieces information in the local profile information in the user's global user profile, such as page identifiers, advertisement identifiers, and rejection indicators. The Robinson-Jaye system additionally provides processes for using the information in the user's global user profile for generating targeted advertising for other users.

Petitioner-Provided Diagram 18: Robinson-Jaye System



73. First, as discussed above, the central system uses the local profile information in the user's global user profile for targeting advertisements to the user. *See supra* ¶¶46-50. When a user visits a page, the Smart Ad Box would appear on the page and the user would see an advertisement targeted to the particular user. LIVEINTENT1005, 4:43-46. To the extent prior uploaded local profile information is available, the central system uses the uploaded local profile information collected from the host machines to determine the advertisement to present to the user in the Smart Ad Box. LIVEINTENT1004, 2:23-51, 3:64-4:6, 4:51-64, 15:14-62, 16:49-18:67; *see* LIVEINTENT1005, 14:22-25.

74. For example, the central system determines from the local profile information interests of the user to identify the advertisement that would be presented in the Smart Ad Box. To determine the user's interests, the central system uses a process that involves: (1) identifying interest categories based on the interest categories of the pages visited by the user, (2) adding these interest categories to the global user profile of the user, and (3) generating scores associated with these interest categories. LIVEINTENT1005, 10:42-46, 11:45-51, 12:17-65; *see* LIVEINTENT1005, 2:50-62, 9:42-12:65, figs.5-6. The central system uses the uploaded local profile information to determine the user interests. LIVEINTENT1004, 15:16-28, LIVEINTENT1005, 10:42-46, 11:45-51; *see* LIVEINTENT1004, 1:27-44, 2:9-22, 15:16-16:67; LIVEINTENT1005, 9:42-

12:65, figs.5-6. And the central system would further use the determined interests of the user to identify targeted advertising to present to the user.

LIVEINTENT1004, 15:16-28; *see* LIVEINTENT1004, 1:27-44, 2:9-22, 2:23-51, 3:64-4:6, 4:51-64, 15:14-18:67; *see also* LIVEINTENT1005, 14:22-25.

75. In more detail, the central system processes the uploaded local profile information to identify the interest categories based on the interest categories of pages visited by the user and adding these interest categories to the global user profile. LIVEINTENT1005, 10:42-46, 11:45-51; *see* LIVEINTENT1005, 9:42-13:18, figs.5-6. The central system determines whether the local profile information contains new interest categories of a site or page visited by the user that were not previously contained in the user's global user profile as part of its process for determining a user's interest profile. *See supra* ¶44.

76. Additionally, the central system computes the score assigned to an interest category for the user by taking into consideration the age of the information about the user so that older information has less impact on an interest score than more recent information. LIVEINTENT1005, 8:55-59; *see* LIVEINTENT1005, 2:58-62. The central system's process for determining an impact of the interest category of the page visited by the user on the calculated score for the user involves scaling a value derived from a content interest score of the interest category of the visited page (variable c in Equation 5 of Table 2) by an

exponential function ($e^{-\left(\frac{\ln(2)}{\gamma}(t_c-t)\right)}$ in Equation 5 of Table 2) that accounts for an age (e.g., a difference between a current time and a time that the user accessed the page) of the information collected about the user's visit to the page.

LIVEINTENT1005, 12:38-41; *see* LIVEINTENT1005, 9:45-10:6, 11:45-12:16, 12:23-42, tbl.1, tbl.2 (equation 5).

77. Second, the central system uses the local user identifiers in the user's global user profile for targeting advertisements to the user. Specifically, the central system uses the local user identifiers for targeting advertisements to the user because the local user identifiers are used for collecting the local profile information that is used for targeting the advertisements. The local user identifiers stored in the global user profile for mapping the local user identifiers to the global user identifier are used by the central system to identify the global user profile to which the local profile information is added, as discussed above. *See supra* ¶¶36-40, 60-68. And the local profile information in the global user profile is used for targeting advertisements. *See supra* ¶¶46-50.

78. Third, the central system uses specific information in the local profile information of the user's global user profile, including, for example, page identifiers, advertisement identifiers, and rejection indicators for selecting advertising to present to the user.

79. The central system uses the page identifier to: (1) determine the subject matter of the page and use the subject matter of the page for selecting an advertisement, and (2) identify whether an advertiser permits an advertisement to be presented on the page visited by the user. Robinson teaches tracking the page identifier of the page visited by the user, and Jaye teaches using page identifiers to track information about the content of pages. LIVEINTENT1004, 9:21-23; LIVEINTENT1005, 5:5-11, 9:42-10:41, fig.5. The central system uses the page identifier of the page including the Smart Ad Box that presents the advertisement to the user in order to identify the page, determine the subject matter of that page, and then select an advertisement that would have “resonance” with and “relate in some way to the subject matter of the page.” LIVEINTENT1004, 4:64-67. Advertisers can also select where their advertisements are displayed by indicating to the central system which sites and pages a particular advertisement should be displayed, and the central system uses the page identifier of the page visited by the user to determine whether an advertiser has identified the visited page as a desirable page or undesirable page for displaying an advertisement. LIVEINTENT1004, 3:29-34, 5:11-21, 5:30-6:15.

80. The central system uses the advertising identifier for selecting an advertisement to present to the user and teaches tracking the identifier of the advertisement that was displayed to the user in the Smart Ad Box.

LIVEINTENT1005, 5:37-40, 5:59-63, 6:4-12; *see* LIVEINTENT1005, 5:37-6:15 (describing the advertisement identifier is used for selecting an advertisement to present on a Smart Ad Box), 9:8-35 (describing that the identifier of the advertisement displayed to the user in the Smart Ad Box is tracked as the user visits a page).

81. As to rejection indicators, when the user rejects an advertisement, the rejected ad would automatically be replaced with another ad targeted to that user. LIVEINTENT1004, 5:4-9. And the central system uses prior rejection indicators as a basis for selecting an advertisement to present to that user. LIVEINTENT1004, 15:37-40; *see* LIVEINTENT1004, 15:15-16:67.

82. In addition to using the user's global user profile to target advertising for the user, the central system uses information in the user's global user profile to target advertising for other users. After the user (referred to as the "first user" in the present discussion) visits the page, the central system uses information in the global user profile about the first user to select an advertisement for a second user. LIVEINTENT1004, 15:16-42; *see* LIVEINTENT1004, 2:31-3:27, 3:61-4:6, 4:9-37, 8:62-10:22, 15:14-18:59. The central system—considering information in global user profiles of a set of other users (including the first user) who are determined to be similar to the second user (LIVEINTENT1004, 15:63-67, 16:53-55, 17:1-3, 17:32-33, 18:44-45)—computes a likelihood that the second user would

be interested in a particular advertisement (LIVEINTENT1004, 15:15-19, 16:3-38, 16:53-67, 17:22-28, 18:18-21, 18:55-59). *See* LIVEINTENT1004, 2:63-67, 16:1-18:59.

83. In more detail, first, to initially identify the set of other users who are similar to the second user, the system computes a degree of similarity or closeness between the second user and the other users by, for example, determining whether the second user and the other users have visited similar or the same tracking-enabled sites, which would correspond to the sources of the information used for, as described next, determining the likelihood that the second user would be interested in the advertisement. LIVEINTENT1004, 15:43-50; *see* LIVEINTENT1004, 2:58-2:62, 15:20-15:62. Second, the central system without human intervention computes the likelihood that the second user would be interested in the particular advertisement based on the tracked information about the other users identified to be similar to the second user, e.g., demographic information, clicks of the advertisement, impressions of the advertisement, and impressions of the advertisement before a click of the advertisement. LIVEINTENT1004, 15:15-19, 16:3-38, 16:53-67, 17:22-28, 18:18-21, 18:55-59; *see* LIVEINTENT1004, 2:63-67, 16:1-18:59.

84. In an example, assume Bob is the first user, Joe is the second user, and the central system is determining whether a particular car-related

advertisement “is likely to be of interest to” Joe. LIVEINTENT1004, 15:15-19.

See LIVEINTENT1004, 1:27-44. Consistent with the discussion of Robinson-Jaye above, first, the central system determines a group of users including Bob “are similar to Joe in their interests,” e.g., because Joe, Bob, and the other users “are remarkably similarly in the sites they’ve chosen to visit” and that would provide certain attributes in the user profiles. LIVEINTENT1004, 15:20-21, 15:46-50.

Second, the central system determines whether the car-related advertisement would be of interest to Joe based on whether selected tracked information (e.g., demographic information, prior clicks or impressions of the advertisement, etc.) in Bob and the other users’ profiles provided by Bob and the other users in their visits to the sites would indicate that Joe would similarly be interested in the car-related advertisement. LIVEINTENT1004, 15:15-19, 15:45-50, 16:3-38, 16:53-67, 17:22-28, 18:18-21, 18:55-59; *see* LIVEINTENT1004, 2:63-67, 16:1-18:59.

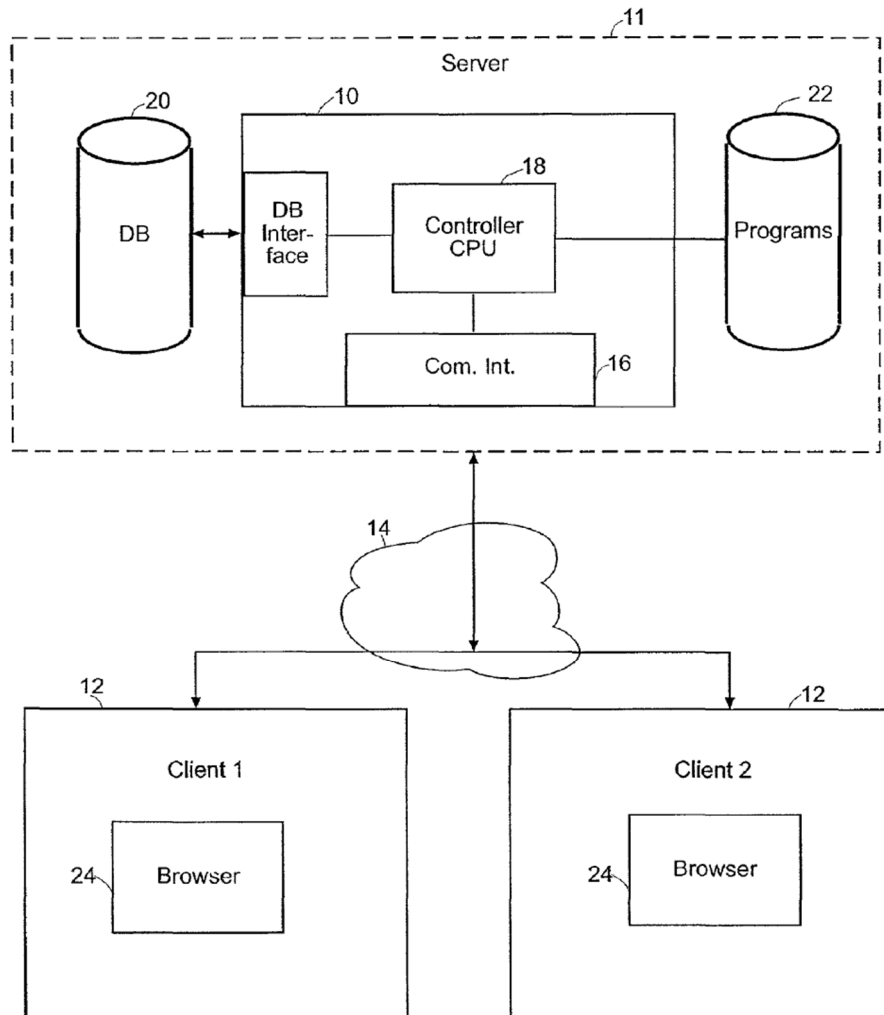
IV. COMBINED TEACHINGS OF ROBINSON, JAYE, AND COLEMAN (ROBINSON-JAYE-COLEMAN)

85. I provide below an analysis of Robinson, Jaye, and U.S. Patent Publication No. 2002/0026351A1 (Coleman) (LIVEINTENT1006) and my opinion of what I believe a POSITA would have found obvious in view the combined teachings of Robinson, Jaye, and Coleman (Robinson-Jaye-Coleman).

A. Coleman Discloses Gathering Information about Users to Create User Profiles, Verifying the Gathered Information, and Selling the Information to Advertisers for Targeted Advertising.

86. Referring to Figure 1 below, Coleman describes a system proprietor that controls operation of an electronic file 11 on a server 10 with which potential purchasers and sellers of goods or services may interact “to form deals or exchange information,” such as advertisements. LIVEINTENT1006 ¶¶[0043], [0047], [0051], [0111], figs.1, 3, 11; *see* LIVEINTENT1006 ¶¶[0053]-[0054], [0058]-[0067], [0079]-[0086], [0091] [0104]-[0117].

Fig. 1



LIVEINTENT1006 fig.1. The electronic file 11 includes profiles of potential purchasers that include information gathered about the potential purchasers.

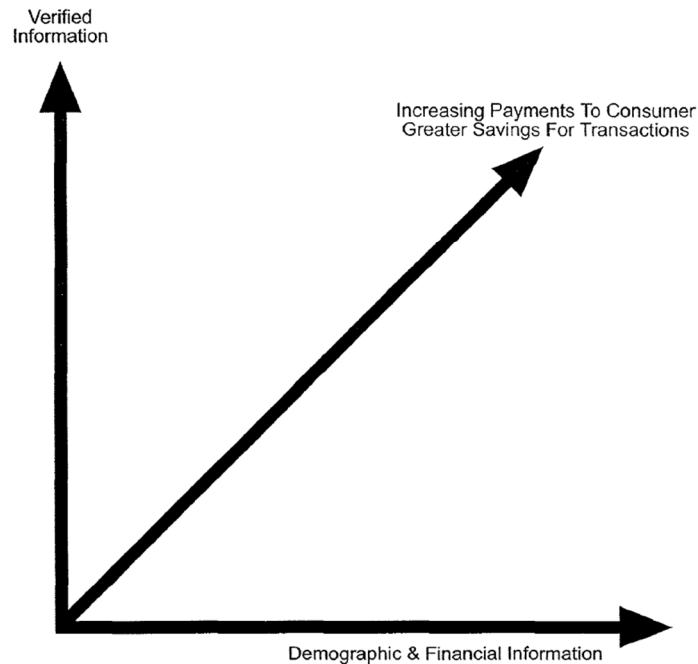
LIVEINTENT1006 ¶¶[0057], [0066], [0104], [0106]; *see* LIVEINTENT1006 ¶¶[0065]-[0077], figs.2-3, 11.

87. The information gathered about the potential purchasers “may be verified for accuracy,” e.g., “by comparing it to information in a variety of other

databases” or by self-verification by a potential purchaser. LIVEINTENT1006

¶¶[0073]-[0074]; *see* LIVEINTENT1006 ¶¶[0049], [0056]. As Coleman describes, referring to Figure 2 below, the value of profile information increases “as the quality or verifiability of the information increases,” as such verification can increase the reliability of predictions that a potential purchaser will buy a good or service offered by a seller. LIVEINTENT1006 ¶[0049], fig.2; *see* LIVEINTENT1006 ¶¶[0073]-[0074]. “After information has been verified, the potential purchaser's profile may be updated to reflect information from which verification with secondary sources was successful and information from which verification from secondary sources was not successful.” LIVEINTENT1006 ¶[0074].

Fig. 2



LIVEINTENT1006 fig.2.

88. Coleman describes that an advertiser can request certain types of information about potential purchasers qualify for the potential purchasers to qualify for the deal that the advertiser wishes to offer. LIVEINTENT1006 ¶¶[0084]-[0086]; *see* LIVEINTENT1006 ¶[0082]. In an embodiment for targeting commercial messages such as advertisements, as shown in Figure 11 below, the server 10 contains profiles 510 of users or potential purchasers and profiles of commercials 512 (e.g., advertisements) that can be “presented over the Internet.” LIVEINTENT1006 ¶[0104]. Advertisers can use the profiles for targeting specific

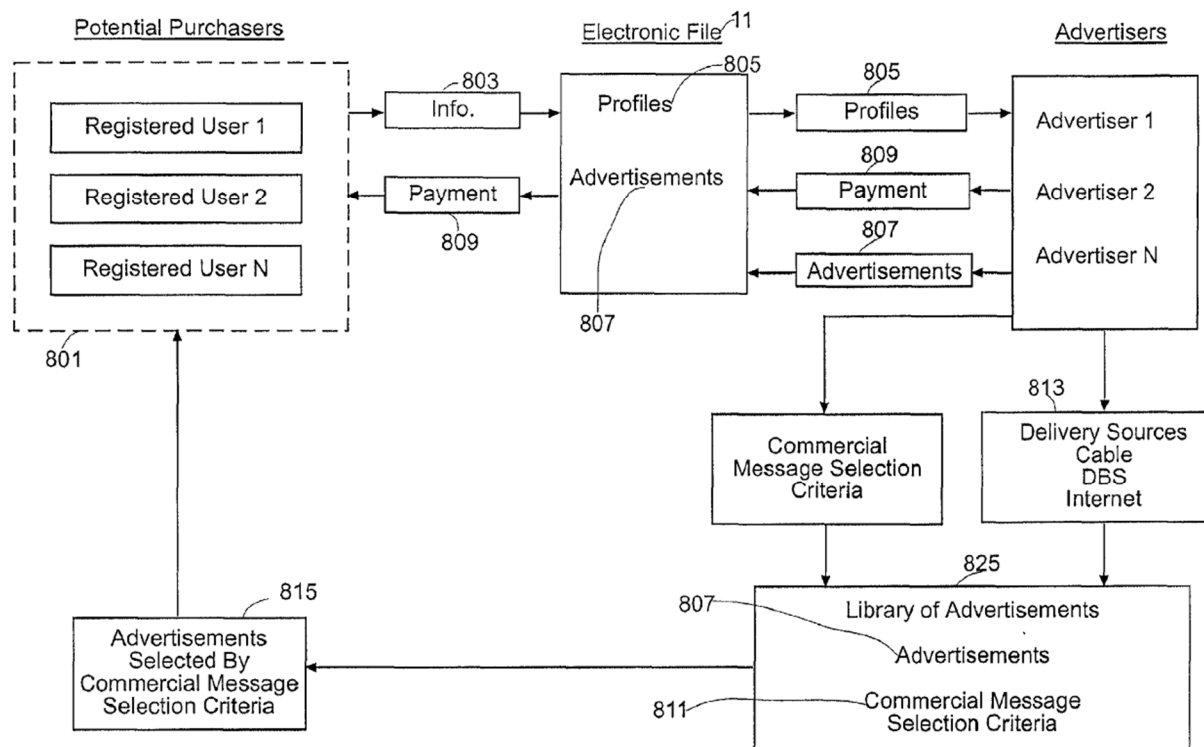
users with advertisements, e.g., that fulfills commercial message selection criteria.

LIVEINTENT1006 ¶¶[0106]-[0107]; *see* LIVEINTENT1006 ¶¶[0106]-[0114].

The advertisers pay the system proprietor (e.g., payment 809 between the electronic file 11 and the advertisers) for the profiles 805 of the users.

LIVEINTENT1006 ¶[0111], fig.11.

Fig. 11



LIVEINTENT1006 fig.11.

B. A POSITA Would Have Been Led to Combine Robinson's, Jaye's, and Coleman's Teachings

89. A POSITA would have found it obvious to implement in Robinson-Jaye an information verification and sales process as taught in Coleman to enable

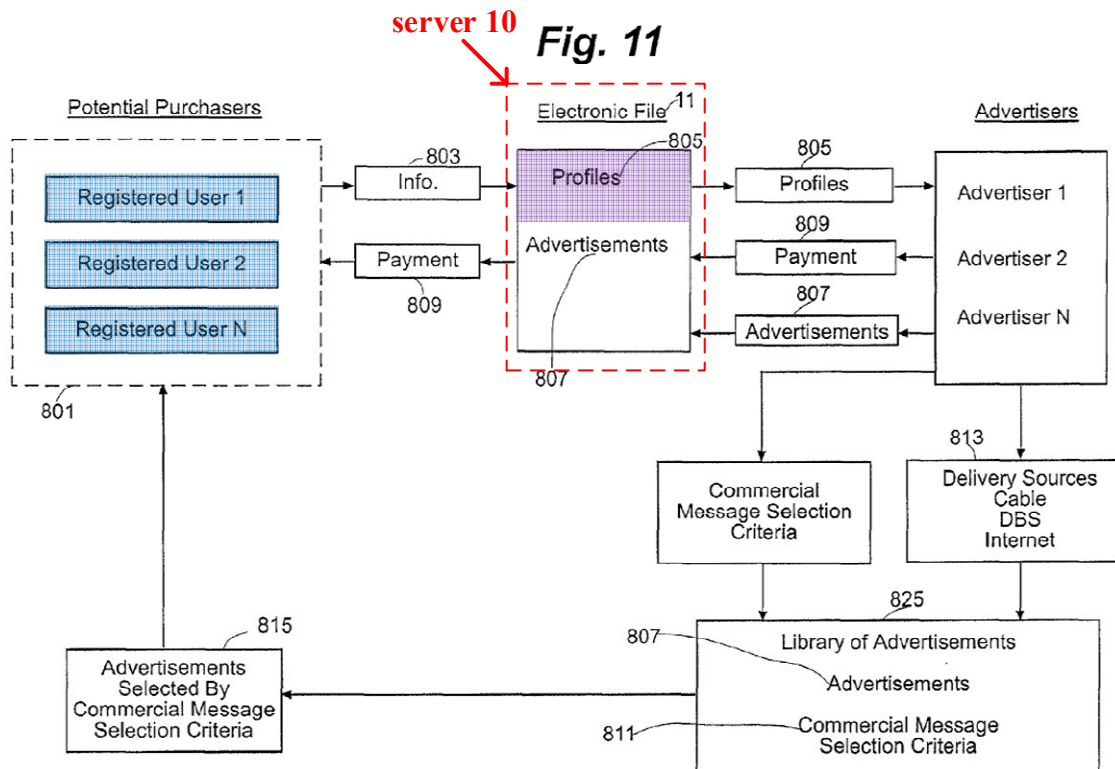
Robinson-Jaye's central system to (1) verify information collected about users in the global user profile, and (2) set a monetary value of the information based on the verification and sell the information.

90. A POSITA would have recognized that Robinson-Jaye is ready for improvement in view of Coleman's teachings. Robinson teaches that participating web sites—analogueous to the owners of the host machines in Robinson-Jaye—pay the company operating the central system for offering the Smart Ad Box and related services for displaying advertisements. LIVEINTENT1004, 13:58-64; *see* LIVEINTENT1004, 4:8-10, 4:43-52. Jaye similarly contemplates that the enterprise server—analogueous to the central system in Robinson-Jaye—releases portions of global user profiles to advertisers. LIVEINTENT1005, 14:22-25.

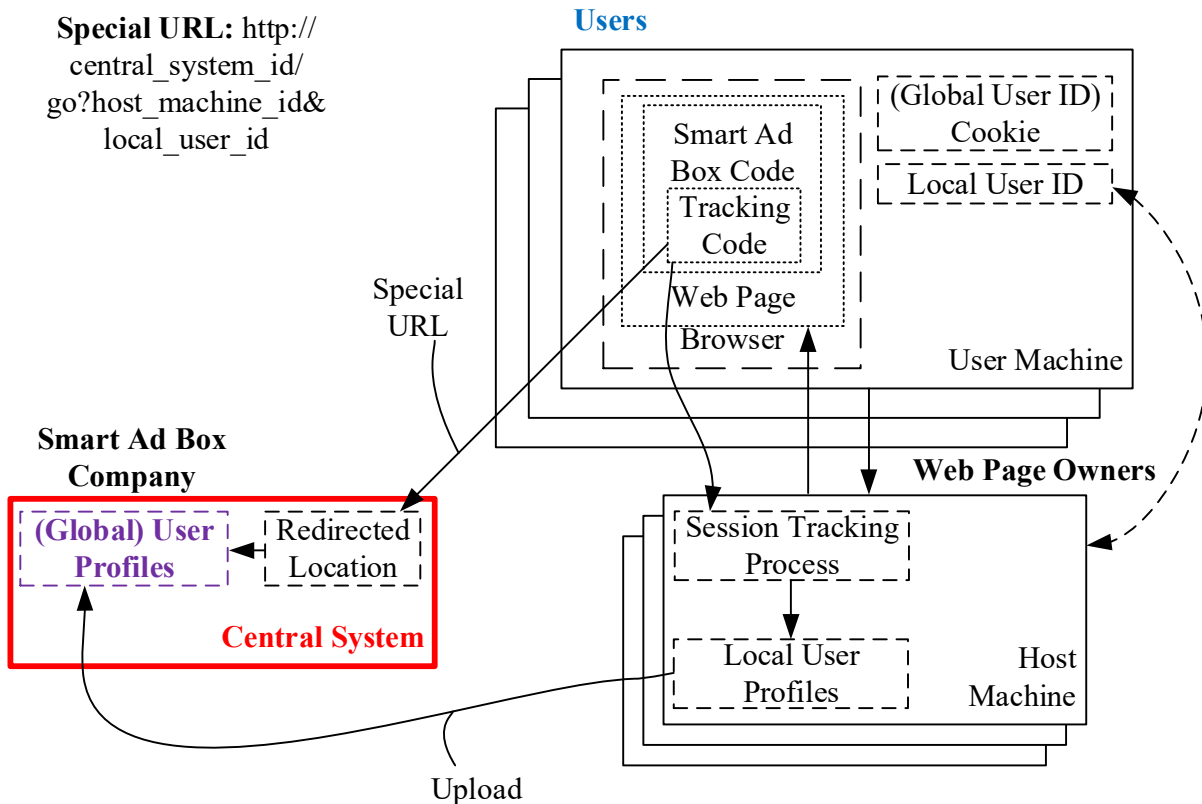
Moreover, Jaye teaches tracking an age of the collected behavior information about users in the global user profiles so that older behavior information has a lesser impact on an interest score for the user. *See* LIVEINTENT1005, 2:58-62, 8:55-59, 9:45-10:6, 11:45-12:42, 12:59-65, tbls.1-2. Based on Coleman's teachings, a POSITA would have recognized a desire to sell specific profile information collected in Robinson-Jaye and set monetary values of the profile information based on the quality or verifiability value of that profile information. *See supra* ¶¶86-88.

91. A POSITA would have found Coleman's teachings to be readily applicable to Robinson-Jaye to yield predictable results. As shown in the diagrams of Robinson-Jaye and Coleman below, like Robinson-Jaye, Coleman teaches a system that involves a **server 10** (analogous to Robinson-Jaye's **central system**) operating the electronic file 11 that stores **profiles** (analogous to the **global user profiles** in Robinson-Jaye) about **users** (analogous to Robinson-Jaye's **users**). LIVEINTENT1006, ¶¶[0055], [0066], [0104], [0106]-[0107], fig.11; *see* LIVEINTENT1006 ¶¶[0106]-[0114]; *supra* §§III.B.2-3. Robinson-Jaye and Coleman similarly teach using information in these user profiles to target advertising to users. LIVEINTENT1006, ¶¶[0104], [0106]-[0114]; *see supra* ¶¶46-50, 86-88. Furthermore, Robinson-Jaye and Coleman both teach quantifying a reliability of this information, with Jaye teaching tracking an age of collected information to determine reliability of the information and Coleman teaching tracking verification of collected information similarly to determine reliability of the information. *See* LIVEINTENT1005, 2:58-62, 8:55-59; LIVEINTENT1006 ¶¶[0049], [0073]-[0074], fig.2; *supra* ¶90. In view of at least these similarities, a POSITA would have recognized that Coleman's teachings would be readily applied to Robinson-Jaye to yield predictable results (i.e., a system that sells information collected about users in the global user profile and that sets a monetary value of the information by verifying the information).

Petitioner-Provided Diagram 19: Annotated Figure 11
of Coleman



Petitioner-Provided Diagram 20: Robinson-Jaye System



92. A POSITA would have been led to implement Coleman's information verification and sales process in Robinson-Jaye such that the central system would (1) verify the information stored in the global user profile and (2) assign a monetary value based on the quality or verifiability of the information and sell the information, as Coleman teaches. LIVEINTENT1006 ¶¶[0049], [0073]; *see supra* ¶¶39-45, 68 (describing types of information tracked in global user profile in Robinson-Jaye). A POSITA would have made such a modification to increase reliability of the information stored in the global user profile, improve advertiser satisfaction in the results of targeting advertisements at users, and improve the

success and consistency of targeted advertising that relies on the information.

LIVEINTENT1006 ¶¶[0049], [0073]-[0074].

93. As to the verification, consistent with Coleman's teachings, a POSITA would have modified Robinson-Jaye to include in the global user profile an indicator of quality or verifiability that indicates whether a piece of information has been successfully verified for accuracy or not. LIVEINTENT1006 ¶¶[0049], [0073]-[0074]. A POSITA would have implemented Robinson-Jaye-Coleman to use various modalities to gather additional information for verifying pieces of information in the global user profile and identifying discrepancies in the pieces of information in the global user profile, e.g., by using public databases and records, third party databases and records, information from self-verification, and information from other information providers. LIVEINTENT1006 ¶¶[0073]-[0075].

94. As to the assigned monetary value and the sale of information, a POSITA would have modified Robinson-Jaye such that the central system sets an increased monetary value for a piece of information in the global user profile if the information is verified, as Coleman teaches. LIVEINTENT1006 ¶[0073]; *see* LIVEINTENT1006 ¶[0049]. The central system would sell to advertisers information in the global user profile to allow the advertisers to target users with advertisements, and the increase in the monetary value of the piece of information

resulting from verification would, as a POSITA would have understood and/or found obvious in view of Coleman's teachings, translate to an increase in the amount that an advertiser would pay for use of the global user profile containing that piece of information. LIVEINTENT1006 ¶¶[0106]-[0107], [0111], fig.11; *see* LIVEINTENT1006 ¶¶[0106]-[0114]. As Coleman teaches, the central system would allow the advertiser to select certain high-value information about users that would allow the advertisers to present advertisements that are likely to elicit user responses. LIVEINTENT1006 ¶¶[0049], [0073], [0085]. And, as a POSITA would have found obvious in view of Coleman's teachings, the advertisers would be able to select particular high-value information (e.g., any of the information discussed above with respect to Robinson-Jaye, including information about a purchase made by the user and demographic and geographic information about the user) that has been verified by certain verification sources. *See* LIVEINTENT1006 ¶¶[0049], [0056], [0073]-[0074], [0084]-[0085]; *supra* ¶¶41-45.

95. Combining Robinson-Jaye's and Coleman's teachings would have merely involved combining known prior art elements (Robinson-Jaye's Smart Ad Box system for tracking and storing user information in a global user profile to present targeted advertising and Coleman's system for verifying and selling user information in a profile) according to known methods (verifying user information in the profile, indicating whether verification of the information was successful,

and setting a monetary value of the information as a function of the verification for sale of that information, as Coleman teaches) to yield predictable results (the central system tracking user information in the global user profile, indicating in the global user profile whether the user information was successfully verified, and setting a monetary value of the information as a function of the verification for sale of that information).

96. A POSITA would have had a reasonable expectation of success combining Robinson-Jaye and Coleman at least because Robinson-Jaye and Coleman disclose similar Internet-enabled systems for tracking information about users. Specifically, as discussed above, Robinson-Jaye and Coleman involve networks including a centralized server (e.g., the central system in Robinson-Jaye and the server 10 in Coleman) that tracks information about users in profiles (e.g., the global user profiles in Robinson-Jaye and the profiles 805 in Coleman) for use in targeted advertising. LIVEINTENT1006, ¶¶[0055], [0066], [0104], [0106]-[0107], fig.11; *see* LIVEINTENT1006 ¶¶[0106]-[0114]; LIVEINTENT1004, 1:59-63, 5:10-13, 9:2-7, 13:61-63; LIVEINTENT1005, 14:22-25; *supra* ¶¶35-50. Moreover, Robinson-Jaye and Coleman both contemplate operation of their systems over the Internet using conventional computer networking technology, including servers, computers, and protocols. LIVEINTENT1006 ¶¶[0043]-[0044];

LIVEINTENT1004, 8:41-50, 8:61-9:20, 9:49-53, 19:27-33; LIVEINTENT1005, 4:5-23, 13:42-62; *see supra* ¶33.

C. An Example of an Implementation of Robinson-Jaye-Coleman

97. I provide below a technical description of an example of an implementation of the Robinson-Jaye-Coleman combination consistent with the discussion above.

98. As discussed above in connection with the example of the implementation of the Robinson-Jaye combination, the host machine generates the local profile information about the user and uploads this information to the central system. *See supra* ¶¶60-68. And the information tracked by the host machine and then uploaded to the central system includes numerous pieces and/or types of information. *See supra* ¶¶41-45, 51-54, 66-71.

99. The central system verifies information in the global user profile and stores an indicator of quality or verifiability of the information in the global user profile. For example, the central system verifies the pieces of information stored in the global user profile. LIVEINTENT1006 ¶¶[0049], [0073]. To verify the pieces of information, the central system uses various modalities to gather additional information for verifying the pieces of information in the global user profile and identifying discrepancies in the pieces of information in the global user profile, e.g., by using public databases and records, third party databases and records,

information from self-verification, and information from other information providers. LIVEINTENT1006 ¶¶[0073]-[0075].

100. The central system then, based on the results of this verification, stores an indicator of quality or verifiability of the information in the global user profile about the user. LIVEINTENT1004, 15:43-56; *see* LIVEINTENT1004, 10:47-64, 15:57-65, 16:37-51; *supra* ¶¶89-93, 95. The indicator of quality or verifiability indicates, for example, whether a piece of information has been successfully verified for accuracy or not. LIVEINTENT1006 ¶¶[0049], [0073]-[0074]; *supra* ¶93.

101. The indicator of quality or verifiability influences a monetary value of the information. The central system assigns a monetary value for the information in the global user profile (e.g., the monetary value that an advertiser would pay to use the information) depending on whether the information is verified (e.g., the monetary value that an advertiser would pay to use the information).

LIVEINTENT1006 ¶[0073]; *see* LIVEINTENT1006 ¶[0049]; *supra* ¶94. For example, the central system sets an increased monetary value for a piece of information in the global user profile if the information is verified.

LIVEINTENT1006 ¶[0073]; *see* LIVEINTENT1006 ¶[0049]; *supra* ¶94.

102. The value paid by advertisers for use of the verified information increases as a function of the increased monetary value resulting from verification.

For example, the central system sells to advertisers information in the global user profile to allow the advertisers to target users with advertisements, and the increase in the monetary value of the piece of information resulting from verification would translate to an increase in the amount that an advertiser would pay for use of the global user profile containing that piece of information. LIVEINTENT1006

¶¶[0106]-[0107], [0111], fig.11; *see* LIVEINTENT1006 ¶¶[0106]-[0114]; *supra* ¶94. The monetary value of a piece of information in the global user profile, and hence the value that an advertiser would pay for the global user profile containing that piece of information, increases if the information is verified.

LIVEINTENT1006 ¶¶[0073], [0106]-[0107], [0111]; *see* LIVEINTENT1006 ¶¶[0049], [0106]-[0114], fig.11; *supra* ¶94.

103. Moreover, advertisers may request information about the user that has been verified by certain sources, and the central system can provide this requested information. For example, the central system allows an advertiser to select certain high-value information about users that would lead to advertisements more likely to elicit user responses, and the central system accordingly selects particular high-value information about the user selected by the advertiser. LIVEINTENT1006 ¶¶[0049], [0073], [0085]; *see supra* ¶94. To make this selection, the advertiser can mandate certain required information about the user before the user is selected for a deal, such as an advertisement, and this required information must have a certain

required threshold value before the advertiser would agree to make the deal with the user. LIVEINTENT1006 ¶¶[0084]-[0085]; *see* LIVEINTENT1006 ¶¶[0055], [0060]-[0061]. The advertiser would be able to mandate particular high-value information that has been verified by certain verification sources and that is therefore more valuable to the advertiser. *See* LIVEINTENT1006 ¶¶[0049], [0056], [0073]-[0074], [0084]-[0085]; *supra* ¶94. The verification sources can include various online sources, credit bureaus, government databases, credit card companies, and any other information providers. LIVEINTENT1006 ¶¶[0074]-[0075]; *see* LIVEINTENT1006 ¶¶[0049], [0056], [0073]-[0074]. Hence, the advertiser is able to request from the central system certain required information that has been verified by certain verification sources. *See* LIVEINTENT1006 ¶¶[0049], [0056], [0073]-[0075], [0084]-[0086]; *supra* ¶94.

V. MANNER IN WHICH THE PRIOR ART REFERENCES RENDER THE '904 PATENT CLAIMS OBVIOUS

104. This declaration explains the conclusions that I have formed based on my analysis. To summarize those conclusions, I believe that:

- a. [GROUND 1A] based upon my knowledge and experience and my review of the prior art publications listed above, claims 1, 3-7, 9-11, 13-17, 19-21, 23-27, and 29-30 are rendered obvious by Robinson-Jaye.

b. **[GROUND 1B]** Based upon my knowledge and experience and my review of the prior art publications listed above, I believe that claims 7-9, 17-19, and 27-29 are rendered obvious by Robinson-Jaye-Coleman.

A. [GROUND 1A] Robinson-Jaye Renders Obvious Claims 1, 3-7, 9-11, 13-17, 19-21, 23-27, and 29-30.

105. A POSITA would have found claims 1, 3-7, 9-11, 13-17, 19-21, 23-27, and 29-30 of the '904 Patent obvious in view of the teachings of Robinson and Jaye, as indicated in the table below.

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
[1.pre]	1. An automated method of collecting profiles of Internet-using entities, the method comprising:	a central system collecting global user profiles of users—including, for each user, local user identifiers, the user's visit to particular pages and sites, information about the user's purchase on the page or site, information about other characteristics of the user—that use user machines to access web pages (<i>see supra</i> ¶¶39-50, 60-75; <i>see, e.g.</i> , LIVEINTENT1004, abstract, 2:13-67, 9:4-37, 10:10-17, 15:29-36, 15:51-54, 18:44-19:5, fig.1; LIVEINTENT1005, abstract, 2:11-3:40, 3:49-60, 6:32-58, 7:46-49, 8:10-9:8, 10:42-46, figs.2, 4-7) The collected profiles contain attributes analogous to attributes disclosed in '904 Patent.

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		LIVEINTENT1001, 2:40-59, 11:21-24, 12:14-15, 12:36-54.
[1.a.i]	(a) electronically receiving at a programmed computer system coupled to a global computer network, from at least one server controlled by one of a plurality of unaffiliated third parties, an electronically URL-redirectioned partial profile of an entity that uses a user computer coupled to the global computer network to access a website,	<i>a programmed computer system coupled to a global computer network:</i> the central system including a central database and a central server (<i>see supra</i> ¶¶56-67; <i>see, e.g.</i>, LIVEINTENT1004, 1:12-15, 1:66-2:8, 5:11-13, 8:40-60, 8:61-9:20, 9:46-10:7, 10:35-44, 19:27-33; LIVEINTENT1005, 1:10-14, 1:15-17, 3:10-14, 3:48-49, 4:5-23, fig.1;) <i>at least one server controlled by one of a plurality of unaffiliated third parties:</i> host machines or servers that host tracking-enabled web pages on participating websites (<i>see supra</i> ¶58; <i>see, e.g.</i>, LIVEINTENT1004, 5:48-50, 6:38-42, 9:17-20, 9:49-50, 13:22-26, 14:1-8; LIVEINTENT1005, 4:46-50, 4:53-5:4, 7:17-26, 8:23-26, 9:42-45, 10:12-21) and that are not owned or controlled by the company that runs the central system (<i>see, e.g.</i>, LIVEINTENT1004, 1:59-63, 4:39-42, 5:10-26, 6:13-37, 8:51-60, 8:64-9:7, 9:12-16, 9:47-53, 13:22-39, 13:58-64) <i>an entity that uses a user computer coupled to the global computer network to access a website:</i> a user that operates a user machine to access a tracking-enabled web page

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		on a participating web site (<i>see supra</i> ¶57; <i>see, e.g.</i>, LIVEINTENT1004, 1:1-3, 1:12-20, 3:61-4:6, 7:52-8:8, 8:21-28, 8:41-67, 9:4-45, 10:23-29; LIVEINTENT1005, 4:7-14, 4:45-5:4) <i>electronically receiving ... an electronically URL-redirected partial profile:</i> the central system, as a result of a forced transfer of the user machine from the host machine to the central system using a special URL, receiving a local user identifier appended to the special URL and receiving local profile information uploaded by a host machine (<i>see supra</i> ¶¶36-40, 60-68; <i>see, e.g.</i>, LIVEINTENT1004, 9:12-16, 9:21-35, 10:8-14; LIVEINTENT1005, 6:16-46, 6:57-7:11, 7:27-36, 7:46-53, 7:58-61, 8:10-43, 10:27-34) The information in the special URL and the uploaded local profile information are analogous to URL-redirected partial profiles disclosed in the '904 Patent. LIVEINTENT1001, 3:2-3, 3:64-4:18, 12:10-24.
[1.a.ii]	which partial profile is available to one of the third parties and contains at least one profile attribute related to the entity,	<i>which partial profile ... contains at least one profile attribute related to the entity:</i> the appended local user identifier and/or the uploaded local profile information include attributes related to the user (<i>see</i>

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		<i>supra</i> §V.A, Elements [1.pre], [1.a.i]) <i>which partial profile is available to one of the third parties:</i> the appended local user identifier and/or the uploaded local profile information are available to the host machine (<i>see supra</i> ¶¶36-40, 60-68; <i>see, e.g.</i>, LIVEINTENT1005, 3:49-50, 5:5-6, 5:31-67, 6:32-46, 7:7-36, 8:35-39, 10:27-32; LIVEINTENT1004, 9:17-20)
[1.a.iii]	which partial profile is received along with an identification of the one of the third parties that contributed the partial profile, and	the appended local user identifier and/or the uploaded local profile information are received along with a host machine identifier of the host machine owned by a third party (<i>see supra</i> §V.A, Element [1.a.i], ¶¶36-40, 51-54, 60-68; <i>see, e.g.</i> , LIVEINTENT1004, 13:58-14:8; LIVEINTENT1005, 6:40-46, 7:46-57)
[1.a.iv]	automatically with the computer system storing the received partial profile;	the central system automatically storing the appended local user identifier and the uploaded local profile information (<i>see supra</i> ¶¶36-40, 51-54, 60-68; <i>see, e.g.</i> , LIVEINTENT1004, 9:4-35, 10:10-14, 15:29-42; LIVEINTENT1005, 6:32-46, 7:21-36, 7:46-49, 8:10-28, 9:33-52, 10:29-36)
[1.b]	(b) automatically with the computer system electronically adding the received partial profile to a maintained profile believed to be related to the same entity;	<i>automatically with the computer system electronically adding the received partial profile to a maintained profile:</i> the central system adding a received partial profile (e.g., the appended local user

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		identifier and/or the uploaded local profile information) to the global user profile (<i>see supra</i> ¶¶36-40, 51-54, 60-68; <i>see, e.g.</i>, LIVEINTENT1004, 9:36-37, 10:10-14, 10:24-26; LIVEINTENT1005, 3:54-55, 6:32-46, 6:57-58, 7:25-45, 7:46-49, 8:10-27, 10:29-41, figs.3-4) <i>a maintained profile believed to be related to the same entity:</i> the global user profile and the received partial profile are related to the same user machine, and hence are believed to be related to the same entity (<i>see supra</i> ¶¶36-40, 51, 63-65; <i>see, e.g.</i>, LIVEINTENT1004, 2:49-57, 8:40-50, 9:36-10:29; LIVEINTENT1005, 5:5-67, 6:6-58, 6:67-8:31, 8:40-52, 9:5-10:41, figs.2-4)
[1.c]	(c) automatically with the computer system generating and storing an electronic record of which of the plurality of unaffiliated third parties contributed to the maintained profile particular profile attributes; and	<i>particular profile attributes:</i> the appended local user identifiers and/or the uploaded local profile information of the user (<i>see supra</i> §V.A, Element [1.pre]) <i>automatically with the computer system generating and storing an electronic record of which of the plurality of unaffiliated third parties contributed to the maintained profile particular profile attributes:</i> the central system associating host machine identifiers with corresponding appended local user

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		identifiers and associating the host machine identifiers with corresponding uploaded local profile information (<i>see supra</i> §V.A, Element [1.a.iii], ¶¶36-42, 51-54, 63-65, 67-69; LIVEINTENT1005, 6:40-46, 7:27-45, 8:10-28)
[1.d]	(d) wherein the maintained profile, including the added partial profile, comprises data used in targeting third-party advertisements to the user computer over the global computer network.	the global user profile, including the received partial profile for the user, includes information (e.g., the local user identifiers and/or the local profile information) that is used to determine the advertisement to present to the user in the Smart Ad Box (<i>see supra</i> §V.A, Element [1.b], ¶¶46-50, 72-84; <i>see, e.g.</i> , LIVEINTENT1004, 1:59-63, 5:10-13, 9:2-7, 13:61-63; LIVEINTENT1005, 14:22-25)
[3]	3. The method of claim 1 wherein the computer system electronically adding the received partial profile to a maintained profile believed to be related to the same entity comprises the computer system adding the received partial profile to a maintained profile believed related to the same entity on account of the received partial profile and the maintained profile being related to the same user computer.	the central system adding the received partial profile to the global user profile on account of the received partial profile and the maintained profile being related to the same user machine (<i>see supra</i> §V.A, Element [1.b], ¶¶36-40, 51, 63-65)
[4]	4. The method of claim 3 wherein the computer system determines that the received partial profile and the maintained profile are related to the same user computer by	<i>the computer system determines that the received partial profile and the maintained profile are related to the same user computer:</i> the central system identifies the global user profile linked to a unique

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
	automatically reading a cookie on the user computer.	global user identifier assigned to the user machine to store the mapping between the local user identifier and the global user identifier and/or the uploaded local profile information (<i>see supra</i> §V.A, Element [1.b], ¶¶36-40, 51, 63-65) <i>by automatically reading a cookie on the user computer:</i> the central system, as a result of the forced transfer, directly receiving from the user machine a cookie containing the global user identifier (<i>see supra</i> §V.A, Element [1.b], ¶¶36-40, 51, 63-65)
[5]	5. The method of claim 1 further comprising automatically with the computer system electronically maintaining an electronic record of which unaffiliated third party contributed, to the maintained profile, the profile attributes used in targeting the advertisements.	the central system storing the host machine identifiers with the <u>local profile information</u> provided by the corresponding host machines (<i>see supra</i> §V.A, Element [1.c], ¶¶39-42, 51-54, 66-69) the central system storing the host machine identifiers with the appended <u>local user identifiers</u> provided by the corresponding host machines (<i>see supra</i> §V.A, Element [1.c], ¶¶36-38, 51-54, 60-65, 69) the central system storing the host machine identifiers with the <u>page identifiers</u> provided by the corresponding host machines (<i>see supra</i> §V.A, Element [1.c], ¶¶39-42, 51-54, 66-69; <i>see, e.g.</i>, LIVEINTENT1004, 3:29-34, 4:64-

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		67, 5:11-21, 5:30-6:15, 9:8-35; LIVEINTENT1005, 5:5-11, 9:42-10:41, fig.5) the central system storing the host machine identifiers with the <u>advertisement identifiers</u> provided by the corresponding host machines (<i>see supra</i> §V.A, Element [1.c], ¶¶39-42, 51-54, 66-69; <i>see, e.g.</i>, LIVEINTENT1004, 5:37-6:15, 5:59-63, 6:4-12, 9:8-35) the central system storing the host machine identifiers with the <u>rejection indicators</u> provided by the corresponding host machines (<i>see supra</i> §V.A, Element [1.c], ¶¶39-42, 51-54, 66-69; <i>see, e.g.</i>, LIVEINTENT1004, 5:4-9, 15:15-16:67) To the extent that that claim 5 requires that the electronic record is distinct from the electronic record of Element [1.c], Robinson-Jaye's electronic record for Element [5] is distinct from Robinson-Jaye's electronic record of Element [1.c]. <i>See supra</i> §V.A, Element [1.c], ¶¶43, 51-54, 70; LIVEINTENT1005, 8:16-28; LIVEINTENT1007, 7:34-54, 9:22-26; LIVEINTENT1001, 7:60-63.
[6]	6. The method of claim 1 further comprising automatically with the computer system electronically determining whether the received	the central system determining whether the uploaded local profile information contains new interest categories of a site or page visited

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
	partial profile contains any profile attributes about the entity that were not previously contained in a maintained profile.	by the user that were not previously contained in the user's global user profile (<i>see supra</i> ¶¶44, 75; <i>see, e.g.</i> , LIVEINTENT1005, 10:42-11:51, 12:17-43, 13:32-36, fig.6)
[7]	7. The method of claim 1 further maintaining with the computer system a credibility rating associated with at least one of the profile attributes of the maintained profile.	the central system indicating an age associated with behavior information in the uploaded local profile information (<i>see supra</i> ¶¶45, 76; <i>see, e.g.</i> , LIVEINTENT1005, 2:58-62, 8:55-59, 9:45-10:6, 11:45-12:42, 12:59-65, tbls.1-2)
[9]	9. The method of claim 1 further comprising automatically with the computer system selecting profile attributes from the maintained profile based on the source thereof.	the central system selecting information from the global user profile of the user (a first user) based on a site that provided the information being similar or the same as a site that provided information for another user (a second user) (<i>see supra</i> ¶¶82-84; <i>see, e.g.</i> , LIVEINTENT1004, 1:27-44, 2:31-3:27, 3:61-4:6, 4:9-37, 8:62-10:22, 15:14-18:59)
[10]	10. The method of claim 1 further comprising automatically with the computer system electronically reading a cookie previously stored on the entity's computer, as a result of the URL redirection.	the central system accessing the cookie—including the global user identifier—that was previously set by the central system and provided to the user machine (<i>see supra</i> §V.A, Elements [1.b], [3], [4], ¶¶36-40, 51, 63-65)
[11]	11. A computer system programmed and connected to perform a method of collecting profiles of Internet-using entities, the method comprising:	<i>A computer system programmed and connected to perform a method of collecting profiles of Internet-using entities:</i> the central system collects global user profiles of users—including, for each user, local user identifiers, the user's visit to particular pages and sites,

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		information about the user's purchase on the page or site, information about other characteristics of the user—that use user machines to access web pages (<i>see supra</i> §V.A, Elements [1.pre], [1.a.i], [1.a.iv])
[11.a.i]	(a) electronically receiving at a computer system coupled to a global computer network, from at least one server controlled by one of a plurality of unaffiliated third parties, an electronically URL-redIRECTED partial profile of an entity that uses a user computer coupled to the global computer network to access a website,	<i>See supra</i> §V.A, Element [1.a.i].
[11.a.ii]	which partial profile is available to one of the third parties and contains at least one profile attribute related to the entity,	<i>See supra</i> §V.A, Element [1.a.ii].
[11.a.iii]	which partial profile is received along with an identification of the one of the third parties that contributed the partial profile, and	<i>See supra</i> §V.A, Element [1.a.iii].
[11.a.iv]	automatically with the computer system storing the received partial profile;	<i>See supra</i> §V.A, Element [1.a.iv].
[11.b]	(b) automatically with the computer system electronically adding the received partial profile to a maintained profile believed to be related to the same entity;	<i>See supra</i> §V.A, Element [1.b].
[11.c]	(c) automatically with the computer system generating and storing an electronic record of which of the plurality of unaffiliated third parties contributed to the maintained	<i>See supra</i> §V.A, Element [1.c]

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
	profile particular profile attributes; and	
[11.d]	(d) wherein the maintained profile, including the added partial profile, comprises data used in targeting third-party advertisements to the user computer over the global computer network.	<i>See supra</i> §V.A, Element [1.d]
[13]	13. The computer system of claim 11 wherein the computer system electronically adding the received partial profile to a maintained profile believed to be related to the same entity comprises the computer system adding the received partial profile to a maintained profile believed related to the same entity on account of the received partial profile and the maintained profile being related to the same user computer.	<i>See supra</i> §V.A, Element [3].
[14]	14. The computer system of claim 13 wherein the computer system determines that the received partial profile and the maintained profile are related to the same user computer by automatically reading a cookie on the user computer.	<i>See supra</i> §V.A, Element [4].
[15]	15. The computer system of claim 11 wherein the method performed by the computer system further comprises automatically with the computer system electronically maintaining an electronic record of which unaffiliated third party contributed, to the maintained profile, the profile attributes used in targeting the advertisements.	<i>See supra</i> §V.A, Element [5].

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
[16]	16. The computer system of claim 11 wherein the method performed by the computer system further comprises automatically with the computer system electronically determining whether the received partial profile contains any profile attributes about the entity that were not previously contained in a maintained profile.	<i>See supra</i> §V.A, Element [6].
[17]	17. The method of claim 1 further maintaining with the computer system a credibility rating associated with at least one of the profile attributes of the maintained profile.	<i>See supra</i> §V.A, Element [7].
[19]	19. The computer system of claim 11 wherein the method performed by the computer system further comprises automatically with the computer system selecting profile attributes from the maintained profile based on the source thereof.	<i>See supra</i> §V.A, Element [9].
[20]	20. The computer system of claim 11 wherein the method performed by the computer system further comprises automatically with the computer system electronically reading a cookie previously stored on the entity's computer, as a result of the URL redirection.	<i>See supra</i> §V.A, Element [10].
[21]	21. A tangible, non-transitory data storage medium comprising indicia of instructions for a processor to perform a method of collecting profiles of Internet-using entities, the method comprising:	<i>A tangible, non-transitory data storage medium comprising indicia of instructions for a processor to perform:</i> software, application programs used in conjunction with such software and/or a computer program tangibly embodied in a

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
		machine-readable storage device for execution by a processor (<i>see, e.g.</i>, LIVEINTENT1004, 4:8-15, 19:27-30; LIVEINTENT1005, 13:42-14:6) <i>a method of collecting profiles of Internet-using entities:</i> a central system collecting global user profiles of users—including, for each user, local user identifiers, the user's visit to particular pages and sites, information about the user's purchase on the page or site, information about other characteristics of the user—that use user machines to access web pages (<i>see supra</i> §V.A, Element [1.pre])
[21.a.i]	(a) electronically receiving at a programmed computer system coupled to a global computer network, from at least one server controlled by one of a plurality of unaffiliated third parties, an electronically URL-redirected partial profile of an entity that uses a user computer coupled to the global computer network to access a website,	<i>See supra</i> §V.A, Element [1.a.i].
[21.a.ii]	which partial profile is available to one of the third parties and contains at least one profile attribute related to the entity,	<i>See supra</i> §V.A, Element [1.a.ii].
[21.a.iii]	which partial profile is received along with an identification of the one of the third parties that contributed the partial profile, and	<i>See supra</i> §V.A, Element [1.a.iii].

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
[21.a.iv]	automatically with the computer system storing the received partial profile;	<i>See supra</i> §V.A, Element [1.a.iv].
[21.b]	(b) automatically with the computer system electronically adding the received partial profile to a maintained profile believed to be related to the same entity;	<i>See supra</i> §V.A, Element [1.b].
[21.c]	(c) automatically with the computer system generating and storing an electronic record of which of the plurality of unaffiliated third parties contributed to the maintained profile particular profile attributes; and	<i>See supra</i> §V.A, Element [1.c]
[21.d]	(d) wherein the maintained profile, including the added partial profile, comprises data used in targeting third-party advertisements to the user computer over the global computer network.	<i>See supra</i> §V.A, Element [1.d]
[23]	23. The data storage medium of claim 21 wherein the computer system electronically adding the received partial profile to a maintained profile believed to be related to the same entity comprises the computer system adding the received partial profile to a maintained profile believed related to the same entity on account of the received partial profile and the maintained profile being related to the same user computer.	<i>See supra</i> §V.A, Element [3].
[24]	24. The data storage medium of claim 23 wherein the computer system determines that the received partial profile and the maintained	<i>See supra</i> §V.A, Element [4].

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
	profile are related to the same user computer by automatically reading a cookie on the user computer.	
[25]	25. The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further comprises: automatically with the computer system electronically maintaining an electronic record of which unaffiliated third party contributed, to the maintained profile, the profile attributes used in targeting the advertisements.	<i>See supra</i> §V.A, Element [5].
[26]	26. The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further comprises: automatically with the computer system electronically determining whether the received partial profile contains any profile attributes about the entity that were not previously contained in a maintained profile.	<i>See supra</i> §V.A, Element [6].
[27]	The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further comprises: maintaining with the computer system a credibility rating associated with at least one of the profile attributes of the maintained profile.	<i>See supra</i> §V.A, Element [7].
[29]	29. The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further	<i>See supra</i> §V.A, Element [9].

Element	Claim Language of '904 Patent	Teachings of Robinson and Jaye
	comprises: automatically with the computer system selecting profile attributes from the maintained profile based on the source thereof.	
[30]	30. The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further comprises: automatically with the computer system electronically reading a cookie previously stored on the entity's computer, as a result of the URL redirection.	<i>See supra</i> §V.A, Element [10].

B. [GROUND 1B] Robinson-Jaye-Coleman Renders Obvious Claims 7-9, 17-19, and 27-29.

106. A POSITA would have found claims 7-9, 17-19, and 27-29 obvious in view of the teachings of Robinson, Jaye, and Coleman, as indicated in the table below.

Element	Claim Language of '904 Patent	Teachings of Robinson, Jaye, and Coleman
[7]	The method of claim 1 further maintaining with the computer system a credibility rating associated with at least one of the profile attributes of the maintained profile.	an indicator of quality or verifiability associated with information in the global user profile (<i>see supra</i> ¶¶89-93, 95, 99-100; <i>see, e.g.</i> , LIVEINTENT1004, 10:47-64, 15:43-65, 16:37-51)
[8]	The method of claim 7 wherein the credibility rating is reflected in a price for the at least one profile attribute.	monetary value of a piece of information in the global user profile increases if the information is verified (<i>see supra supra</i> ¶¶89-92, 94-95, 101-102; <i>see, e.g.</i> , LIVEINTENT1006 ¶¶[0049], [0073], [0106]-[0114], fig.11;)

Element	Claim Language of '904 Patent	Teachings of Robinson, Jaye, and Coleman
[9]	The method of claim 1 further comprising automatically with the computer system selecting profile attributes from the maintained profile based on the source thereof.	the computer system automatically selecting particular high-value information verified by certain verification sources (<i>see supra</i> ¶¶41-45, 89-95, 103; <i>see, e.g.</i> , LIVEINTENT1006 ¶¶[0049], [0055]-[0056], [0060]-[0061], [0073]-[0075], [0084]-[0086])
[17]	The computer system of claim 11 wherein the method performed by the computer system further comprises maintaining with the computer system a credibility rating associated with at least one of the profile attributes of the maintained profile.	<i>See supra</i> §V.B, Element [7].
[18]	The computer system of claim 17 wherein the credibility rating is reflected in a price for the at least one profile attribute.	<i>See supra</i> §V.B, Element [8].
[19]	The computer system of claim 11 wherein the method performed by the computer system further comprises automatically with the computer system selecting profile attributes from the maintained profile based on the source thereof.	<i>See supra</i> §V.B, Element [9].
[27]	The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further comprises: maintaining with the computer system a credibility rating associated with at least one of the profile attributes of the maintained profile.	<i>See supra</i> §V.B, Element [7].

Element	Claim Language of '904 Patent	Teachings of Robinson, Jaye, and Coleman
[28]	The data storage medium of claim 27 wherein the credibility rating is reflected in a price for the at least one profile attribute.	<i>See supra</i> §V.B, Element [8].
[29]	The data storage medium of claim 21 wherein the indicia of instructions instruct the processor to perform a method that further comprises: automatically with the computer system selecting profile attributes from the maintained profile based on the source thereof.	<i>See supra</i> §V.B, Element [9].

VI. LEVEL OF ORDINARY SKILL IN THE ART

107. My analysis and conclusions set forth in this Declaration are based on the perspective of a person of ordinary skill in the art having the level of knowledge and skill as of the invention date of the '904 Patent (POSITA). Based on instruction from Counsel (Fish & Richardson P.C.), I have applied the earliest possible priority date of the '904 Patent—December 13, 1999—as the invention date (Critical Date).

108. I was asked to consider the patent claims and the prior art through the eyes of a POSITA at the Critical Date. I understand that the factors considered in determining the ordinary level of skill in a field of art include the level of education and experience of persons working in the field; the types of problems encountered in the field; the teachings of the prior art, and the sophistication of the

technology at the Critical Date. I understand that a POSITA is not a specific real individual, but rather is a hypothetical individual having the qualities reflected by the factors above. I understand that a POSITA would also have knowledge from the teachings of the prior art, including the art cited in this Declaration.

109. Taking these factors into consideration, based on my knowledge and experience in the field and my review of the '904 Patent and file history, I believe that a POSITA in this matter would have had at least a combination of (1) at least a bachelor's degree in computer science, computer engineering, electrical engineering, or a related field, and (2) at least two years of industry experience with e-commerce computing systems, including systems that support targeted advertising. Additional education could substitute for professional experience, and professional experience could substitute for education. *Id.*

110. Based on my experiences, I have a good understanding of the capabilities of a POSITA. Indeed, I have taught, participated in organizations with, and worked closely with many such persons over the course of my career. Based on my knowledge, skill, and experience, I have an understanding of the capabilities of a POSITA. For example, from my industry experience, I am familiar with what a POSITA would have known and found predictable in the art. From teaching and supervising my undergraduate, graduate, and post-graduate students, I also have an understanding of the knowledge that a person with this academic experience

possesses, and, more generally, people with various levels of academic experience possess. Furthermore, I possess those capabilities myself.

VII. LEGAL PRINCIPLES

A. Terminology

111. I understand that claim terms are generally given their plain and ordinary meaning in light of the patent's specification and file history as understood by a POSITA as of the invention date—here, the Critical Date. In that regard, I understand that the best indicator of claim meaning is its usage in the context of the patent specification as understood by a POSITA. I further understand that the words of the claims should be given their plain meaning unless that meaning is inconsistent with the patent specification or the patent's history of examination before the U.S. Patent and Trademark Office (Office). I also understand that the words of the claims should be interpreted as they would have been interpreted by a POSITA at the Critical Date.

B. Legal Standards

112. I have been informed by Counsel and understand that documents and materials that qualify as prior art can render a patent claim unpatentable as being anticipated or obvious.

113. I have been informed by Counsel and understand that all prior art references are to be looked at from the viewpoint of a person of ordinary skill in

the art at the time of the invention, and that this viewpoint prevents one from using his or her own insight or hindsight in deciding whether a claim is anticipated or rendered obvious.

C. Anticipation

114. I understand that patents or printed publications that qualify as prior art can be used to invalidate a patent claim as anticipated or as obvious.

115. I understand that, once the claims of a patent have been properly construed, the second step in determining anticipation of a patent claim requires a comparison of the properly construed claim language to the prior art on a limitation-by-limitation basis.

116. I understand that a prior art reference “anticipates” an asserted claim, and thus renders the claim invalid, if all limitations of the claim are disclosed in that prior art reference, either explicitly or inherently (i.e., necessarily present).

D. Obviousness

117. I understand that even if a patent is not anticipated, it is still invalid if the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a POSITA.

118. I have been informed by Counsel and understand that a claim is unpatentable for obviousness and that obviousness may be based upon a

combination of references. I have been informed by Counsel and understand that the combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results. However, I have been informed by Counsel and understand that a patent claim composed of several elements is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art.

119. I have been informed by Counsel and understand that when a patented invention is a combination of known elements, a court determines whether there was an apparent reason to combine the known elements in the fashion claimed by the patent at issue by considering the teachings of prior art references, the effects of demands known to people working in the field or present in the marketplace, and the background knowledge possessed by a person having ordinary skill in the art.

120. I have been informed by Counsel and understand that a patent claim composed of several limitations is not proved obvious merely by demonstrating that each of its limitations was independently known in the prior art. I am informed by Counsel and understand that identifying a reason those elements would be combined can be important because inventions in many instances rely upon building blocks long since uncovered, and claimed discoveries almost of necessity will be combinations of what, in some sense, is already known. I am informed by

Counsel and understand that it is improper to use hindsight in an obviousness analysis, and that a patent's claims should not be used as a "roadmap."

121. I have been informed by Counsel and understand that an obviousness inquiry requires consideration of the following factors: (1) the scope and content of the prior art, (2) the differences between the prior art and the claims, (3) the level of ordinary skill in the art, and (4) any so called "secondary considerations" of non-obviousness, which include: (i) "long felt need" for the claimed invention, (ii) commercial success attributable to the claimed invention, (iii) unexpected results of the claimed invention, and (iv) "copying" of the claimed invention by others.

122. I have been informed by Counsel and understand that an obviousness evaluation can be based on a single reference or a combination of multiple prior art references. I understand that the prior art references themselves may provide a suggestion, motivation, or reason to combine, but that the nexus linking two or more prior art references is sometimes simple common sense. I have been informed by Counsel and understand that obviousness analysis recognizes that market demand, rather than scientific literature, often drives innovation, and that a motivation to combine references may be supplied by the direction of the marketplace.

123. I have been informed by Counsel and understand that if a technique has been used to improve one device, and a person of ordinary skill at the time of

invention would have recognized that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill.

124. I have been informed by Counsel and understand that practical and common sense considerations should guide a proper obviousness analysis, because familiar items may have obvious uses beyond their primary purposes. I have been informed by Counsel and understand that a person of ordinary skill looking to overcome a problem will often be able to fit together the teachings of multiple prior art references. I have been informed by Counsel and understand that obviousness analysis therefore takes into account the inferences and creative steps that a person of ordinary skill would have employed at the time of invention.

125. I have been informed by Counsel and understand that a proper obviousness analysis focuses on what was known or obvious to a person of ordinary skill at the time of invention, not just the patentee. Accordingly, I understand that any need or problem known in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining the elements in the manner claimed.

126. I have been informed by Counsel and understand that a claim can be obvious in light of a single reference, without the need to combine references, if

the elements of the claim that are not found explicitly or inherently in the reference can be supplied by the common sense of one of skill in the art.

127. I have been informed by Counsel and understand that there must be a relationship between any such secondary considerations and the invention, and that contemporaneous and independent invention by others is a secondary consideration supporting an obviousness determination.

128. In sum, my understanding is that prior art teachings are properly combined where one of ordinary skill having the understanding and knowledge reflected in the prior art and motivated by the general problem facing the inventor, would have been led to make the combination of elements recited in the claims. Under this analysis, the prior art references themselves, or any need or problem known in the field of endeavor at the time of the invention, can provide a reason for combining the elements of multiple prior art references in the claimed manner.

129. I have been informed by Counsel and understand that in an IPR, “the petitioner shall have the burden of proving a proposition of unpatentability,” including a proposition of obviousness, “by a preponderance of the evidence.” 35 U.S.C. §316(e).

VIII. CONCLUSION

130. I am over the age of 18 and am competent to write this declaration. I have personal knowledge, or have developed knowledge of these technologies based upon education, training, or experience, of the matters set forth herein.

131. I am not and never have been, an employee of LiveIntent. I received no compensation for this declaration beyond my normal hourly compensation based on my time actually spent analyzing the '904 Patent, the prior art publications cited below, and issues related thereto, and I will not receive any added compensation based on the outcome of any IPR or other proceeding involving the '904 Patent.

132. In writing this Declaration, I have considered the following: my own knowledge and experience, including my work experience in the fields of networking, algorithms and data structures, electronic commerce, artificial intelligence, computer systems, and databases; my experience in teaching those subjects; and my experience in working with others involved in those fields.

133. My opinions, as explained above, are based on my education, experience, and expertise in the fields relating to the '904 Patent. Any figures that appear within this document have been prepared with the assistance of Counsel and reflect my understanding of the '904 Patent and the prior art discussed above.

In addition, I have analyzed the following publications and materials, in addition to other materials I refer to or cite in my declaration:

- a. Robinson;
- b. Jaye; and
- c. Coleman.
- d. the '904 Patent and excerpts from its prosecution history
(LIVEINTENT1002).

134. I conclude from this analysis, as discussed in the reasoning noted in this declaration, 1, 3-11, 13-21, and 23-30 of the '904 Patent are obvious.

135. I currently hold the opinions set expressed in this declaration. But my analysis may continue, and I may acquire additional information and/or attain supplemental insights that may result in added observations.

APPENDIX A

Michael Mitzenmacher
michaelm@eecs.harvard.edu
33 Cary Avenue, Lexington MA 02421
617-496-7172

Research Interests	Design and Analysis of Algorithms; Machine Learning Systems; Networks and Data Transmission; Databases.
Education	UNIVERSITY OF CALIFORNIA AT BERKELEY, Berkeley, CA Ph.D. in Computer Science awarded December, 1996. Dissertation: The Power of Two Choices in Randomized Load Balancing. Advisor: Alistair Sinclair. GPA: 4.0/4.0 CAMBRIDGE UNIVERSITY, Cambridge, England Attended as one of ten recipients of the Churchill Fellowship. Cambridge C.A.S. in Mathematics with highest distinction awarded June 1992. HARVARD COLLEGE, Cambridge, MA B.A. in Mathematics with Computer Science, summa cum laude, awarded June 1991.
Employment	HARVARD UNIVERSITY, Cambridge, MA <i>Spring 1999-present</i> Assistant professor (from Jan. 1999 -July 2002), Associate professor (from July 2002-January 2005), Professor (from Jan. 2005-present), Area Dean for Computer Science (from July 2010-June 2013), Area Co-Chair for Computer Science (from July 2018-June 2019). Teach undergraduate courses including “Introduction to Algorithms and Data Structures” and “Discrete Mathematics for Computer Science”, and graduate courses covering topics in randomized algorithms, algorithms for networks, compression, coding, cryptography, and information retrieval. DIGITAL SYSTEMS RESEARCH CENTER, Palo Alto, CA <i>Fall 1996-Winter 1998</i> Research scientist. Projects included work on information retrieval on the Web, erasure codes, error-correcting codes, on-line algorithms, and load balancing. Co-inventor for twelve submitted patents. SANTA CLARA UNIVERSITY, Santa Clara, CA <i>Spring 1997</i> Guest professor for the undergraduate class “Introduction to Algorithms.” Consultant: I consult on intellectual property issues as an expert witness and in other capacities. As an expert witness, I have testified in multiple trials. I have also consulted for several technology companies and research laboratories, including Adverplex (Cogolabs), Akamai, AT&T, Digital Fountain, eHarmony, Fluent Mobile (Fiksu), Google, Huawei, ITA Software, JobSync, Microsoft, Mitsubishi Research Laboratories, Weta Digital, and Yahoo.
Funding	NSF CCF-2101140: AF: Small: Algorithms and Data Structures with Predictions. PI: Michael Mitzenmacher. Total grant: \$400,000. 7/21-6/24. NSF CNS-2107078: CNS Core: Medium: Approximation and Randomization in the Programmable Data Plane. co-PIs: Minlan Yu. Total grant: \$1,200,000. 7/21-6/25.

NSF DMS-2023528: Foundations of Data Science Institute (Harvard University). 9/20-8/25.

NSF CCF-1563710: CIF: NeTS: Medium: Collaborative Research: Unifying Data Synchronization. co-PIs: David Starobinski, Ari Trachtenberg. Total grant: \$400,000. 7/16-6/20.

NSF CCF-1535705: AitF: FULL: Collaborative Research: Better Hashing for Applications: From Nuts & Bolts to Asymptotics. co-PIs: David Andersen. Total grant: \$250,000. 9/15-8/19.

NSF CCF-1320231: AF: Small: Data Synchronization : Theory, Algorithms, and Practice PI: Michael Mitzenmacher. Total grant: \$399,370. 9/13-8/16.

NSF CNS-1228598: TWC: Medium: Collaborative: Privacy-Preserving Distributed Storage and Computation. co-PIs: Michael Goodrich, Roberto Tamassia. Total grant: \$400,000. 8/12-8/16.

NSF IIS-0964473: HCC: Medium: Collaborative Research: Data-Parallel Hash Tables: Theory, Practice and Applications. co-PIs: Nina Amenta. Total grant: \$171,095. 8/10-7/13.

NSF CCF-0915922: AF : Small : The Theory and Practice of Hash-Based Algorithms and Data Structures. PI: Michael Mitzenmacher. Total grant: \$441,956. 8/09-7/12.

NSF CNS-0721491: NeTS FIND: A Network-Wide Hashing Infrastructure for Monitoring and Measurement. PI: Michael Mitzenmacher. Total grant: \$330,000. 9/07-8/10.

NSF CCF-0634923: Towards a Basic Understanding of Channels with Synchronization Errors. PI: Michael Mitzenmacher. Total grant: \$200,000. 9/06-8/09.

NSF CCR-0121154: ITR/SY Algorithmic Issues in Large Scale Dynamic Networks. co-PIs: Eli Upfal, Brown. Subcontract to Harvard. Total grant: \$502,000. 9/01-8/06.

NSF CCR-0118701: Low Density Parity Check Codes for Channels with Memory. co-PIs: Alek Kavcic. Total grant: \$510,000. 9/01-8/04.

NSF CCR-9983832: Dynamic Processes and Network Algorithms (CAREER). \$200,000. 7/00-6/04.

Dean's Competitive Fund for Promising Scholarship. 2018. \$49,333.

Google University Research Program. 8/13-8/14. PIs: Eddie Kohler and Michael Mitzenmacher \$56,000.

Google University Research Program. 12/11-12/12. \$25,000.

Yahoo! University Research Program. 7/11-6/12. \$10,000.

Google University Research Program. 12/09-12/10. \$60,000.

Yahoo! University Research Program. 9/09-8/10. \$10,000.

Google University Research Program. 8/08-7/09. \$75,000.

Yahoo! University Research Program. 9/07-8/08. \$25,000.

Yahoo! University Research Program. 9/06-8/07. \$50,000.

Cisco University Research Program. 8/08-7/09. \$80,000.

Cisco University Research Program. 8/07-7/08. \$83,000.

Cisco University Research Program. 8/06-7/07. \$80,000.

Cisco University Research Program. 8/05-7/06. \$72,000.

Alfred P. Sloan Research Fellowship. \$40,000. Awarded in 2000.

IBM Faculty Research Grant. \$10,000 Awarded in 2005.

IBM Faculty Research Grant. \$10,000 Awarded in 2003.

Mitsubishi Electronic Research Laboratory. \$10,000 for undergraduate research projects. Awarded in 2002.

Honors

Applied Cryptography and Network Security Conference Test of Time Paper Award (2025)

ACM Paris Kanellakis Theory and Practice Award (2020)

IEEE Fellow (2020)

IEEE Senior Member (2018)

ACM Fellow (2014)

ACM Symposium on Parallelism in Algorithms and Architectures Best Paper Award (2014)

World Wide Web Conference Best Paper Award (2014)

Royal Academy of Engineering Distinguished Visiting Fellowship (2010)

ACM SIGCOMM Test of Time Paper Award (2009)

IEEE Information Theory Society Best Paper Award (2002)

Alfred P. Sloan Research Fellowship (2000)

Sakrisson Award (for Ph.D. thesis at Berkeley) (1997)

NDSEG Graduate Fellowship (1992-95)

NSF Graduate Fellowship Winner (1991)

Churchill Fellowship (1991-92)

Hoopess Prize (for senior thesis at Harvard) (1991)

Phi Beta Kappa (1990)

Harvard Distinction in Teaching Award (1990)

Goldwater Fellowship (1989-91)

Professional Activities

Editorships and Boards:

Advisory Board, Hariri Institute at Boston University (2018-2021)

Editorial Board, Leibniz International Proceedings in Informatics (2009-2019)

Editorial Board, Communications of the ACM (2013-2018)

Science Board, Santa Fe Institute (2013-2019)

Guest Editor, SIAM Journal of Computing special issue for STOC 2009

SIAM Journal on Computing, Editor (2006-2010)

Internet Mathematics, Managing Editor

Guest Editor, Theory of Computing System special issue for SPAA 2002

Journal of Interconnection Networks, Editor

Organizing Committees:

Sublinear Algorithms Workshop Organizing Committee, FODSI (2022)
 Algorithms with Predictions Workshop Organizing Committee, STOC (2022)
 Machine Learning for Algorithms Workshop Organizing Committee, FODSI (2021)
 SIGACT Chair (2015-2018)
 Dagstuhl Seminar: Theory and Applications of Hashing (2017)
 EADS Summer School on Hashing: Theory and Applications (2014)
 ICERM Workshop on Stochastic Graph Models (2014)
 WAW 2013 Organizing Committee
 SIGACT Vice-Chair (2009-2012)
 General Chair, STOC (2010)
 Workshop on Randomized Algorithms and Random Graphs (2010)
 DARPA Information Science and Technology Study Groups (2007-2008)
 SIGACT Committee for the Advancement of Theoretical Computer Science (2005-2008)
 MSRI Workshop on Models of Real-World Random Networks (2005)
 FOCS 2003 Local Arrangements Chair
 Second Workshop on Randomized Algorithms and Random Graphs (2003)
 Workshop on Algorithms and Models for the Web Graph (2002)
 First Workshop on Randomized Algorithms and Random Graphs (2002)
 DIMACS Workshop on Quality of Service Issues in the Internet (2001)
 BU/NSF Workshop on Internet Measurement, Instrumentation, and Characterization (1999)
 FOCS 1998 Local Arrangements Chair

Program Committees:

NSDI 2025 Program Committee
 HotNets 2024 Program Committee
 WADS 2023 Program Committee
 ACM CoNEXT 2023 Program Committee
 NeurIPS 2022 Program Committee
 ACM CoNEXT 2022 Program Committee
 ICML 2022 Program Committee
 SIGCOMM 2022 Program Committee
 NeurIPS 2021 Program Committee
 ICML 2021 Program Committee
 ACDA 2021 Program Committee
 ICLR 2021 Program Committee
 APoCS 2021 Program Committee
 ACM CoNEXT 2020 Program Committee
 NeurIPS 2020 Program Committee
 ICML 2020 Program Committee
 APoCS 2020 Program Committee
 NeurIPS 2019 Program Committee
 SIGCOMM 2019 Program Committee
 SOSA 2019 Program Committee (Co-Chair)
 ACM CoNEXT 2018 Program Committee
 ANALCO 2018 Program Committee
 SIGCOMM 2017 Program Committee
 NSDI 2017 Program Committee
 ACM CoNEXT 2016 Program Committee
 SIGCOMM 2016 Program Committee
 ALENEX 2016 Program Committee (Co-Chair)
 ICALP 2016 Program Committee (Track C PC Chair)

ACM CoNEXT 2015 Program Committee
ICALP 2015 Program Committee
ITCS 2015 Program Committee
SIGCOMM 2014 Program Committee
ACM CoNEXT 2014 Program Committee
WSDM 2014 Senior Program Committee
WAW 2013 Program Committee (Co-Chair)
STOC 2013 Executive Committee
WSDM 2013 Senior Program Committee
SIGCOMM 2012 Program Committee
ISIT 2012 Program Committee
NSDI 2012 Program Committee
WAW 2011 Program Committee
ACM CoNEXT 2010 Program Committee
SIGCOMM 2010 Program Committee
NSDI 2010 Program Committee
FUN 2010 Program Committee
LATIN 2010 Program Committee
NetSciCom 2009 Program Committee
SIGCOMM 2009 Program Committee
STOC 2009 Program Committee (PC Chair)
NSDI 2009 Program Committee
ICALP 2008 Program Committee
SPAA 2008 Program Committee
WAW 2007 Program Committee
Analysis of Algorithms 2007 Program Committee
ANALCO 2007 Program Committee
ICALP 2007 Program Committee
RANDOM 2007 Program Committee
NCA 2006 Program Committee
STOC 2006 Program Committee
PODC 2005 Program Committee
ANALCO 2005 Program Committee
PODC 2004 Program Committee
SIGCOMM 2004 Program Committee
IPTPS 2004 Program Committee
Data Compression Conference 2004 Program Committee
ESA 2004 Program Committee
WAW 2003 Program Committee
FOCS 2003 Program Committee
Data Compression Conference 2003 Program Committee
ESA 2002 Program Committee
PODC 2002 Program Committee
SPAA 2002 Program Committee
STOC 2002 Program Committee
Data Compression Conference 2002 Program Committee
ALENEX 2002 Program Committee
HiPC 2001 Program Committee
ISAAC 2000 Program Committee
RANDOM 2000 Program Committee
STACS 2000 Program Committee

FOCS 99 Program Committee
RANDOM 98 Program Committee

Reviewer for several conferences, journals, and grant panels.

Books

Probability and Computing: Randomized Algorithms and Probabilistic Analysis, by Michael Mitzenmacher and Eli Upfal. Published by Cambridge University Press in 2005. (2nd edition available in 2017.) This is a textbook meant for an advanced undergraduate or beginning graduate class. The textbook has been used in courses at Brown, Harvard, U. C. Berkeley, Univ. of Victoria, Tufts, Univ. of Mass. at Amherst, Purdue, U. Penn., and several other universities.

**Conference
and Journal
Publications**

“Don’t Stop Me Now: Embedding Based Scheduling For LLMs,” with R. Shahout, E. Malach, C. Liu, W. Jiang, and M. Yu. To appear in *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025.

“Parallel Peeling of Invertible Bloom Lookup Tables in a Constant Number of Rounds,” with M. Goodrich and R. Kitagawa. To appear in *Proceedings of the 50th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM)*, 2025.

“Efficient d -ary Cuckoo Hashing at High Load Factors by Bubbling Up,” with W. Kuszmaul. To appear in *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, 2025.

“Leveraging Parameterized Chernoff Bounds for Simplified Algorithm Analyses,” with M. Dillencourt and M. Goodrich. *Information Processing Letters*, 187:106516, 2025.

“Optimal and Approximate Adaptive Stochastic Quantization,” with R. Basat, Y. Ben-Itzhak, and S. Vargaftik. In *Proceedings of the 38th Conference on Neural Information Processing Systems*, 2024.

“SkipPredict: When to Invest in Predictions for Scheduling,” with R. Shahout. In *Proceedings of the 38th Conference on Neural Information Processing Systems*, 2024.

“Beyond Throughput and Compression Ratios: Towards High End-to-end Utility of Gradient Compression,” with W. Han, S. Vargaftik, B. Karp, and R. Basat. In *Proceedings of the Twenty-Third ACM Workshop on Hot Topics in Networks*, pp. 186-194, 2024.

“Learning-Augmented Frequency Estimation in Sliding Windows,” with R. Shahout and I. Sabek. In *Proceedings of the Workshop on Intelligent Classification of High-Speed Network Traffic (at ICNP ’24)*.

“Learning-Based Heavy Hitters and Flow Frequency Estimation in Streams,” with R. Shahout. In *Proceedings of the 32nd IEEE International Conference on Network Protocols*, 2024.

“Tabula: Efficiently Computing Nonlinear Activation Functions for Secure Neural Network Inference,” with M. Lam, V. J. Reddi, G. Wei, and D. Brooks. In *Transactions on Machine Learning Research*, 2024.

“Accelerating Federated Learning with Quick Distributed Mean Estimation,” with S. Vargaftik, R. Basat, A. Portnoy, G. Mendelson, and Y. Ben-Itzhak. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 3410-3442, 2024.

- “THC: Accelerating Distributed Deep Learning Using Tensor Homomorphic Compression,” with M. Li, R. Basat, S. Vargaftik, C. Lao, K. Xu, X. Tang, and, M. Yu. In *Proceedings of NSDI*, pp. 1191-1211, 2024.
- “Separating k -Player from t -Player One-Way Communication, with Applications to Data Streams,” with E. Du, D. Woodruff, and G. Yang. In *Theory of Computing*, vol. 19, Article 10, pp. 1-44, 2023.
- “Direct Telemetry Access,” with J. Langlet, R. Basat, G. Oliaro, M. Yu, and G. Antichi. In *Proceedings of ACM SIGCOMM*, pp. 832-849, 2023.
- “Can Learned Models Replace Hash Functions?” with I. Sabek, K. Vaidya, D. Horn, A. Kipf, and T. Kraska. *Proceedings of the VLDB Endowment*, vol. 16, no. 3, pp. 532-545, 2022.
- “Algorithmic Tools for Mining the Motif Structure of Networks,” with B. Matejek, T. Chen, and C. Tsourakakis. To appear in *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2022.
- “Edge-colored directed subgraph enumeration on the connectome,” with B. Matejek, D. Wei, T. Chen, C. Tsourakakis, and H. Pfister. *Scientific Reports*, 2022.
- “Algorithms with Predictions,” with S. Vassilvitskii. *Communications of the ACM*, Volume 65 Issue 7, pp. 33-35, 2022.
- “EDEN: Communication-Efficient and Robust Distributed Mean Estimation for Federated Learning,” with S. Vargaftik, R. Basat, A. Portnoy, G. Mendelson, and Y. Ben-Itzhak. In *Proceedings of the 39th International Conference on Machine Learning*, pp. 21984-22014, 2022.
- “SNARF: A Learning-Enhanced Range Filter,” with K. Vaidya, S. Chatterjee, E. Knorr, S. Idreos, and T. Kraska. In *Proceedings of the VLDB Endowment*, 15(8):1632-1664, 2022.
- “Proteus: A Self-Designing Range Filter,” with E. Knorr, B. Lemaire, A. Lim, S. Luo, H. Zhang, and S. Idreos. In *Proceedings of the SIGMOD International Conference on Management of Data*, pp. 1670-1684, 2022.
- “Designing Sensitive Viral Diagnostics with Machine Learning,” with H. Metsky, N. Welch, P. Pillai, N. Haradhvala, L. Rumker, S. Mantena, Y. Zhang, D. Yang, C. Ackerman, J. Weller, P. Blainey, C. Myhrvold, and P. Sabeti. *Nature Biotechnology*, 2022.
- “The Supermarket Model with Known and Predicted Service Times,” with M. Dell’Amico. *IEEE Transactions on Parallel and Distributed Systems*, 2022.
- “Uniform Bounds for Scheduling with Job Size Estimates,” with Z. Scully and I. Grosz. In *Proceedings of 13th Innovations in Theoretical Computer Science Conference*, 114:1-114:30, 2022.
- “DRIVE: One-bit Distributed Mean Estimation,” with S. Vargaftik, R. Basat, A. Portnoy, G. Mendelson, and Y. Ben-Itzhak. In *Proceedings of the 34th Conference on Neural Information Processing Systems*, pp. 362-377, 2021.
- “Zero-CPU Collection with Direct Telemetry Access,” with J. Langlet, R. Basat, S.S. Ramanathan, G. Oliaro, M. Yu, and G. Antichi. In *Proceedings of the Twentieth ACM Workshop on Hot Topics in Networks*, pp. 108-115, 2021.

“Gradient Disaggregation: Breaking Privacy in Federated Learning by Reconstructing the User Participant Matrix,” with M. Lam, V. J. Reddi, D. Brooks, and G. Wei. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 5959-5968, 2021.

“Putting the ‘Learning’ into Learning-Augmented Algorithms for Frequency Estimation,” with E. Du and F. Wang. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 2860-2869, 2021.

“Queues with Small Advice.” In *Proceedings of the SIAM Conference on Applied and Computational Discrete Algorithms*, pp. 1-12, 2021.

“How to Send a Real Number Using a Single Bit (and Some Shared Randomness),” with R. Basat and S. Vargaftik. In *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming*, 25:1-25:20, 2021.

“SALSA: Self-Adjusting Lean Streaming Analytics,” with R. Basat, G. Einzinger, and S. Vargaftik. In *The 37th IEEE International Conference on Data Engineering*, pp. 864-875, 2021.

“Partitioned Learned Bloom Filters,” with K. Vaidya, E. Knorr, and T. Kraska. In *The Ninth International Conference on Learning Representations (ICLR)*, 2021.

“Improved Sublinear Time Algorithm for Longest Increasing Subsequence,” with S. Seddighin. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, pp. 1934-1947, 2021.

“DISCOVering the heavy hitters with disaggregated sketches,” with V. Bruschi, R. Basat, Z. Liu, G. Antichi, and G. Bianchi. In *Proceedings of the 16th International Conference on Emerging Networks Experiments and Technologies (CoNEXT)*, pp. 536-537, 2020.

“Detecting Routing Loops in the Data Plane,” with J. Kucera, R. Basat, M. Kuka, G. Antichi, and M. Yu. In *Proceedings of the 16th International Conference on Emerging Networks Experiments and Technologies (CoNEXT)*, pp. 466-473, 2020.

“Optimal Learning of Joint Alignments with a Faulty Oracle,” with K. Larsen and C. Tsourakakis. In *Proceedings of ISIT*, pp.2492-2497, 2020.

“PINT: Probabilistic In-band Network Telemetry,” with R. Basat, S. Ramanathan, Y. Li, G. Antichi, and M. Yu. In *Proceedings of ACM SIGCOMM*, pp. 662-680, 2020.

“Dynamic Algorithms for LIS and Distance to Monotonicity,” with S. Seddighin. In *Proceedings of the 52nd Annual Symposium on Theory of Computing*, pp. 671-684, 2020.

“Clustering with a Faulty Oracle,” with K. Larsen and C. Tsourakakis. In *Proceedings of the Web Conference*, pp. 2831-2834, 2020.

“Prophets, Secretaries, and Maximizing the Probability of Choosing the Best,” with H. Esfandiari, M. Hajiaghayi, and B. Lucier. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, pp. 3717-3727, 2020.

“Faster and More Accurate Measurement through Additive-Error Counters,” with R. Basat, G. Einzinger, and S. Vargaftik. In *Proceedings of IEEE INFOCOM*, pp. 1251-1260, 2020. o “Scheduling with Predictions and the Price of Misprediction.” In *Proceedings of 11th Innovations in Theoretical Computer Science Conference*, 14:1-14:18, 2020.

“Equitability, Interval Estimation, and Statistical Power,” with Y. Reshef, D. Reshef, and P. Sabeti. *Statistical Science*, (35)2:202-217, 2020.

“A Digital Fountain Retrospective,” with J. Byers and M. Luby. *ACM SIGCOMM Computer Communication Review*, 49(5), pp. 82-85, 2019.

“Robust Set Reconciliation via Locality Sensitive Hashing,” with T. Morgan. In *Proceedings of the 38th Symposium on Principles of Database Systems*, pp. 164-181, 2019.

“Online Pandora’s Boxes and Bandits,” with H. Esfandiari, M. Hajiaghayi, and B. Lucier. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pp. 1885-1892, 2019.

“Arithmetic Progression Hypergraphs: Examining the Second Moment Method.” In *Proceedings of ANALCO*, pp. 127-134, 2019.

“A Bayesian Nonparametric View on Count-Min Sketch,” with R. Adams and D. Cai. In *Proceedings of the 31st Conference on Neural Information Processing Systems*, pp. 8782-8791, 2018.

“A Model for Learned Bloom Filters and Optimizing by Sandwiching.” In *Proceedings of the 31st Conference on Neural Information Processing Systems*, pp. 462-471, 2018.

“Metric Sublinear Algorithms via Linear Sampling,” with H. Esfandiari. In *Proceedings of the 59th Annual IEEE Symposium on Foundations of Computer Science*, pp. 11-22, 2018.

“EMOMA: Exact Match in One Memory Access,” with S. Pontarelli and P. Reviriego. *IEEE Transactions on Knowledge and Data Engineering*, 30(11):2120-2133, 2018.

“Weightless: Lossy Weight Encoding for Deep Neural Network Compression,” with B. Reagen, U. Gupta, R. Adolf, A. Rush, G. Wei, and D. Brooks. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 4321-4330, 2018.

“Load Thresholds for Cuckoo Hashing with Double Hashing,” with K. Panagiotou and S. Walzer. In *Proceedings of the 16th Scandinavian Symposium and Workshops on Algorithm Theory*, pp. 29:1-29:9, 2018.

“Joint Alignment From Pairwise Differences with a Noisy Oracle,” with C. Tsourakakis. In *Proceedings of the 15th Workshop on Algorithms and Models for the Web Graph*, pp. 59-69, 2018. Journal version: *Internet Mathematics*, at <https://doi.org/10.24166/im.06.2019>, December 2020.

“Reconciling Graphs and Sets of Sets,” with T. Morgan. In *Proceedings of the 37th Symposium on Principles of Database Systems*, pp. 33-47, 2018.

“Simulated Annealing for JPEG Quantization,” with M. Hopkins and S. Wagner-Carena. In *Proceedings of the 2018 IEEE Data Compression Conference*, p. 414.

“Adaptive Cuckoo Filters,” with S. Pontarelli and P. Reviriego. In *Proceedings of the ALENEX Conference*, pp. 36-47, 2018. Journal version: *ACM Journal of Experimental Algorithmics*, Article 1.1, March 2020.

“An Empirical Comparison of the Maximal and Total Information Coefficients to Leading Measures of Dependence,” with D. Reshef, Y. Reshef, and P. Sabeti. *Annals of Applied Statistics*, 12(1):123-155, 2018.

- “Simple Multi-Party Set Reconciliation,” with R. Pagh. *Distributed Computing* 31(6):441-453, 2018.
- “Technical Perspective: Building a Better Hash Function.” *Communications of the ACM*, 60(7), p. 93, 2017.
- “Compresso: Efficient Compression of Segmentation Data For Connectomics,” with D. Haehn, F. Lekschas, B. Matejek, and H. Pfister. In *Proceedings of the 20th International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 781-788, 2017.
- “Auditable Data Structures,” with M. Goodrich, E. Kornaropoulos, and R. Tamassia. In *Proceedings of the European Symposium on Security and Privacy*, pp. 285-300, 2017.
- “2-3 Cuckoo Filters for Faster Triangle Listing and Set Intersection,” with D. Eppstein, M. Goodrich, and M. Torres. In *Proceedings of the 36th Symposium on Principles of Database Systems*, pp. 247-260, 2017.
- “Scalable motif-aware graph clustering,” with C. Tsourakakis and J. Pachocki. In *Proceedings of the 26th International Conference on the World Wide Web*, pp. 1451-1460, 2017.
- “Measuring Dependence Powerfully and Equitably,” with Y. Reshef, D. Reshef, H. Finucane, and P. Sabeti. *Journal of Machine Learning Research*, 17(212):1-63, 2016.
- “Quantized Random Projections and Non-Linear Estimation of Cosine Similarity,” with P. Li and M. Slawski. In *Advances In Neural Information Processing Systems 29*, pp.2748-2756, 2016.
- “Analyzing Distributed Join-Idle-Queue: A Fluid Limit Approach.” In *Proceedings of the 54th Annual Allerton Conference on Communication, Control, and Computing*, pp. 312-318, 2016.
- “Hardness of Peeling with Stashes,” with V. Nathan. *Information Processing Letters*, 116(11), pp. 682-688, 2016.
- “More Practical and Secure History-Independent Hash Tables,” with M. Goodrich, E. Kornaropoulos, and R. Tamassia. In *European Symposium on Research in Computer Security*, pp. 20-38, 2016.
- “Models and Algorithms for Graph Watermarking,” with D. Eppstein, M. Goodrich, J. Lam, N. Mamano, and M. Torres. In *Proceedings of the Information Security Conference*, pp. 283-301, 2016.
- “Better Bounds for Coalescing-Branching Random Walks on Graphs,” with R. Rajaraman and S. Roche. In *Proceedings of the 28th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pp. 313-323, 2016. Journal version: *ACM Transactions on Parallel Computing*, 5(1), 2:1-2:23, 2018.
- “OMASS: One Memory Access Set Separation,” with S. Pontarelli and P. Reviriego. *IEEE Transactions on Knowledge and Data Engineering*, 28(7):1940-1943, 2016.
- “Voronoi Choice Games”, with M. Boppana, R. Hod, and T. Morgan. In *Proceedings of 43rd International Colloquium on Automata, Languages, and Programming*, 23:1-23:13, 2016.

“Space Lower Bounds for Itemset Frequency Sketches,” with E. Liberty, J. Thaler, and J. Ullman. In *Proceedings of the 35th Symposium on Principles of Database Systems*, pp. 441-454, 2016.

“Technical Perspective: Catching Lies (and Mistakes) in Offloaded Computation,” with J. Thaler. *Communications of the ACM*, 59(2), p. 102, 2016.

“A New Approach to Analyzing Robin Hood Hashing.” In *Proceedings of ANALCO*, pp. 10-24, 2016.

“More Analysis of Double Hashing for Balanced Allocations.” In *Proceedings of ANALCO*, pp. 1-9, 2016.

“Theory Without Experiments: Have We Gone Too Far?” *Communications of the ACM*, 58(9), pp. 40-42, 2015.

“Scaling Up Clustered Network Appliances with XBricks,” with D. Zhou, B. Fan, H. Lim, R. Wang, M. Kaminsky, D. Andersen, and A. Singh. In *Proceedings of ACM SIGCOMM*, 2015.

“Scalable Large Near-Clique Detection in Large-Scale Networks via Sampling,” with J. Pachocki, R. Peng, C. Tsourakakis, and S. Chen Xu. In *Proceedings of the 21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2015.

“Repeated Deletion Channels,” with B. Haeupler. In *Proceedings of the IEEE Information Theory Workshop*, pp. 152-156, 2014.

“Multi-Party Set Reconciliation Using Characteristic Polynomials,” with A. Boral. In *Proceedings of the 52nd Annual Allerton Conference on Communication, Control, and Computing*, pp. 1182-1187, 2014.

“Cleaning Up the Record on the Maximal Information Coefficient and Equitability,” with D. Reshef, Y. Reshef, and P. Sabeti. *Proceedings of the National Academy of Sciences* 111(33):E3362-3, 2014.

“Cuckoo Filter: Practically Better than Bloom,” with B. Fan, D.G. Andersen, and M. Kaminsky. In *Proceedings of the 10th International Conference on Emerging Networks Experiments and Technologies (CoNEXT)*, pp. 75-88, 2014.

“Coding for Random Projections,” with P. Li and A. Shrivastava. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, pp. 676-684, 2014.

“Balanced Allocations and Double Hashing.” In *Proceedings of the 26th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pp. 331-342, 2014.

“Parallel Peeling Algorithms,” with J. Jiang and J. Thaler. In *Proceedings of the 26th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pp. 319-330, 2014. Journal version: *ACM Transactions on Parallel Computing*, 3(1), 7, 2016.

“Wear Minimization for Cuckoo Hashing: How Not to Throw a Lot of Eggs into One Basket”, with D. Eppstein, M. Goodrich, and P. Pszona. In *Proceedings of the 13th International Symposium on Experimental Algorithms*, pp. 162-173, 2014.

“Efficient Estimation for High Similarities using Odd Sketches,” with R. Pagh and N. Pham. In *Proceedings of the 23rd International Conference on the World Wide Web*, pp. 109-118, 2014.

“Improving the Performance of Invertible Bloom Lookup Tables,” with S. Pontarelli and P. Reviriego. *Information Processing Letters*, 114(4), pp. 185-191, 2014.

“Special Section on the Forty-First Annual ACM Symposium on Theory of Computing (STOC 2009),” with N. Immerlica, J. Katz, R. Servedio, and C. Umans. *SIAM Journal on Computing*, 41(6):1591-1592, 2012.

“The Daily Deals Marketplace: Empirical Observations and Managerial Implications,” with J. Byers and G. Zervas. *AC SIGecom Exchanges*, 11(2):29-31, 2012.

“Peeling Arguments and Double Hashing,” with J. Thaler. In *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1118-1125, 2012.

“The Complexity of Object Reconciliation, and Open Problems Related to Set Difference and Coding,” with G. Varghese. In *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1126-1132, 2012.

“Cache-Oblivious Dictionaries and Multimaps with Negligible Failure Probability,” with M. Goodrich, D. Hirschberg, and J. Thaler. In *Proceedings of the First Mediterranean Conference on Algorithms*, pp. 203-218, 2012.

“An Economic Analysis of User-Privacy Options in Ad-Supported Services,” with J. Feigenbaum and G. Zervas. In *Proceedings of the 8th Workshop on Internet and Network Economics*, pp. 30-43, 2012.

“Verifiable Computation with Massively Parallel Interactive Proofs,” with H. Pfister, M. Roberts, and J. Thaler. In *Proceedings of the USENIX HotCloud Workshop, 2012*.

“Biff (Bloom Filter) Codes: Fast Error Correction for Large Data Sets,” with G. Varghese. In *Proceedings of the International Symposium on Information Theory*, pp. 483-487, 2012.

“Continuous Time Channels with Interference,” with I. Ivan, J. Thaler, and H. Yuen. In *Proceedings of the International Symposium on Information Theory*, pp. 860-864, 2012.

“The Groupon Effect on Yelp Ratings: a Root Cause Analysis,” with J. Byers and G. Zervas. In *Proceedings of the ACM Conference on Electronic Commerce*, pp. 248-265, 2012.

“Anonymous Card Shuffling and Its Applications to Parallel Mixnets,” with M. Goodrich. In *Proceedings of 39th International Colloquium on Automata, Languages, and Programming*, pp. 549-560, 2012.

“Chernoff-Hoeffding Bounds for Markov Chains: Generalized and Simplified,” with K. Chung, H. Lam, and Z. Liu. In *Proceedings of the 29th International Symposium on Theoretical Aspects of Computer Science*, pp. 124-135, 2012.

“Practical oblivious storage,” with M. Goodrich, O. Ohrimenko and R. Tamassia. In *Proceedings of the Second ACM Conference on Data and Application Security and Privacy*, pp. 13-24, 2012.

- “Hierarchical Heavy Hitters with the Space Saving Algorithm,” T. Steinke and J. Thaler. In *Proceedings of 2012 Meeting on Algorithm Engineering and Experiments*, pp. 160-174.
- “Daily Deals: Prediction, Social Diffusion, and Reputational Ramifications,” with J. Byers and G. Zervas. In *Proceedings of the 5th International Conference on Web Search and Web Data Mining*, pp. 543-552, 2012.
- “Practical Verified Computation with Streaming Interactive Proofs,” with G. Cormode and J. Thaler. In *Proceedings of Innovations in Theoretical Computer Science*, pp. 90-112, 2012.
- “Information Dissemination via Random Walks in d -Dimensional Space,” with H. Lam, Z. Liu, X. Sun, and Y. Wang. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1612-1622, 2012.
- “Privacy-Preserving Group Data Access via Stateless Oblivious RAM Simulation,” with M. Goodrich and R. Tamassia. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 157-167, 2012.
- “An Efficient Rigorous Approach for Identifying Statistically Significant Frequent Itemsets,” with A. Kirsch, A. Pietracaprina, G. Pucci, E. Upfal, and F. Vandin. *Journal of the ACM*, 59(3), 2012.
- “Detecting Novel Associations in Large Datasets,” with D. Reshef, Y. Reshef, H. Finucane, S. Grossman, G. McVean, P. Turnbaugh, E. Lander, and P. Sabeti. *Science*, Vol. 3004, No. 6602, pp. 1518-1524, December 2011.
- “External-Memory Multimaps,” with E. Angelino, M. Goodrich, and J. Thaler. In *Proceedings of the 22nd International Symposium on Algorithms and Computation*, pp. 384-394, 2011. Journal version: *Algorithmica*, 67(1), pp. 23-48, 2013.
- “Brief announcement: large-scale multimaps,” with M. Goodrich. In *Proceedings of the 23rd ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pp. 259-260, 2011.
- “Oblivious RAM Simulation with Efficient Worst-Case Access Overhead,” with M. Goodrich, O. Ohrimenko, and R. Tamassia. In *Proceedings of 3rd ACM Cloud Computing Workshop*, pp. 95-100.
- “Invertible Bloom Lookup Tables,” with M. Goodrich. In *Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing*, pp. 792-799, 2011.
- “Cuckoo Hashing with Pages,” with M. Dietzfelbinger and M. Rink. In *Proceedings of the 19th Annual European Symposium on Algorithms*, pp. 615-627, 2011.
- “Network Coding Meets TCP: Theory and Implementation,” with J.K. Sundararajan, D. Shah, M. Medard, S. Jakubczak, and J. Barros. *Proceedings of the IEEE*, 99(3), pp. 490-510, 2011.
- “Privacy-Preserving Access of Outsourced Data via Oblivious RAM Simulation,” with M. Goodrich. In *Proceedings of 38th International Colloquium on Automata, Languages, and Programming*, pp. 576-587, 2011.

- “Graption: A Graph-based P2P Traffic Classification Framework for the Internet Backbone,” with M. Iliofotou, H. Kim, M. Faloutsos, P. Pappu, and G. Varghese. *Computer Networks*, 55(8), pp. 1909-1920, 2011.
- “On the Zero-Error Capacity Threshold for Deletion Channels,” with I. Kash, J. Thaler, J. Ullman. In *Proceedings of the Information Theory and Applications Workshop*, pp. 1-5, 2011.
- “Heapable Sequences and Subsequences,” with J. Byers, B. Heeringa, and G. Zervas. In *Proceedings of ANALCO*, pp. 33-44, 2011.
- “An Introduction to Human Guided Search.” *ACM Crossroads*, 17(2):34-35, 2010.
- “Popularity is Everything: A New Approach to Protecting Passwords from Statistical Guessing Attacks,” with S. Schechter and C. Herley. In *Proceedings of the 5th Usenix Workshop on Hot Topics in Security*, pp. 1-8, 2010.
- “Streaming Graph Computations with a Helpful Advisor,” with G. Cormode and J. Thaler. In *Proceedings of the European Symposium on Algorithms*, pp. 231-242, 2010. Journal version: *Algorithmica*, 65(2), pp. 409-442, 2013.
- “Tight Thresholds for Cuckoo Hashing via XORSAT,” with M. Dietzfelbinger, A. Goerdt, A. Montanari, R. Pagh, and M. Rink. In *Proceedings of 37th International Colloquium on Automata, Languages, and Programming*, pp. 213-225, 2010.
- “Local cluster aggregation models of explosive percolation,” with R.M. D’Souza. *Physical Review Letters*, 14(19), p. 104-107, 2010.
- “Tight Asymptotic Bounds for the Deletion Channel with Small Deletion Probabilities,” with A. Kalai and M. Sudan. In *Proceedings of the International Symposium on Information Theory*, pp. 997-1001, 2010.
- “An Improved Analysis of the Lossy Difference Aggregator,” with H. Finucane. *Computer Communication Review*, 40(2), pp. 4-11, 2010.
- “Information Asymmetries in Pay-Per-Bid Auctions: How Swoopo Makes Bank,” with J. Byers and G. Zervas. In *Proceedings of the ACM Conference on Electronic Commerce*, pp. 1-12, 2010.
- “Carousel: Scalable Logging for Intrusion Prevention Systems”, with T. Lam and G. Varghese. In *Proceedings of NSDI 2010*, pp. 361-376.
- “AMS Without 4-Wise Independence on Product Domains,” with V. Braverman, K. Chung, Z. Liu, and R. Ostrovsky. In *Proceedings of the 27th International Symposium on Theoretical Aspects of Computer Science*, pp. 119-130, 2010.
- “Adaptive Weighing Designs for Keyword Value Computation,” with J. Byers and G. Zervas. In *Proceedings of the Third International Conference on Web Search and Web Data Mining* pp. 331-340, 2010.
- “Human-guided search,” with G. Klau, N. Lesh, and J. Marks. *Journal of Heuristics*, 16(3), pp. 289-310, 2010.

“Real-Time Parallel Hashing on the GPU,” with D. Alcantara, A. Sharf, F. Abbasinejad, S. Sengupta, J. Owens, and N. Amenta. *ACM Transactions on Graphics*, 28(5), p. 154, 2009.

“Exploiting Dynamicity in Graph-based Traffic Analysis: Techniques and Applications,” with M. Iliofotou and M. Faloutsos. In *Proceedings of ACM Co-NEXT*, pp. 241-252, 2009.

“An Analysis of Random-Walk Cuckoo Hashing,” with A. Frieze and P. Mellsted. In *Proceedings of 2009 APPROX-RANDOM*, pp. 490-503, 2009. Journal version: *SIAM Journal on Computing*, 40(2), pp. 291-308, 2011.

“Some Open Questions Related to Cuckoo Hashing.” In *Proceedings of the 17th Annual European Symposium on Algorithms*, pp. 1-10, 2009.

“On Compressing Social Networks,” with F. Chierichetti, R. Kumar, S. Lattanzi, A. Panconesi, and P. Raghavan. In *Proceedings of the 15th ACM SIGKDD Int’l Conference on Knowledge Discovery and Data Mining*, pp. 219-228, 2009.

“An Efficient Rigorous Approach for Identifying Statistically Significant Frequent Itemsets,” with A. Kirsch, A. Pietracaprina, G. Pucci, E. Upfal, and F. Vandin. In *Proceedings of the 28th Symposium on Principles of Database Systems*, pp. 117-126, 2009.

“Using the Power of Two Choices to Improve Bloom Filters,” with Steven Lumetta. *Internet Mathematics*, vol. 4. no. 1, pp. 17-33, 2009.

“Network coding meets TCP,” with J. Sundararajan, D. Shah, M. Medard, and J. Barros. In *Proceedings of IEEE INFOCOM*, pp. 280-288, 2009.

“An Economically Principled Generative Model of AS Graph Connectivity,” with J. Corbo, S. Jain, and D. Parkes. In *Proceedings of IEEE INFOCOM (Mini-conference)*, pp. 2941-2945, 2009.

“Designing floating codes for expected performance,” with H. Finucane and Z. Liu. In *Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing*, 2008. Journal version (with F. Chierichetti, H. Finucane, and Z. Liu): *IEEE Transactions on Information Theory*, vol 56, Issue 3, pp. 968-978, 2010.

“On the performance of multiple choice hash tables with moves on deletions and inserts,” with A. Kirsch. In *Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1284-1290, 2008.

“More Robust Hashing: Cuckoo Hashing with a Stash,” with A. Kirsch and U. Wieder. In *Proceedings of the 16th Annual European Symposium on Algorithms*, pp. 611-622, 2008. Journal version: *SIAM Journal on Computing*, 39(4), pp. 1543-1561, 2009.

“A survey of results for deletion channels and related synchronization channels.” In *Proceedings of the 2008 Scandinavian Workshop on Algorithm Theory*, pp. 1-3, 2008. Journal version: “New Results and Open Problem for Channels with Synchronization.” *Probability Surveys*, 2009.

“The Power of One Move: Hashing Schemes for Hardware,” with A. Kirsch. In *Proceedings of IEEE INFOCOM*, pp. 106-110, 2008. Journal version: *IEEE/ACM Transactions on Networking*, 18(6):1752-1765, 2010.

“Distributed Beamforming with Binary Signaling,” with M. Johnson and K. Ramchandran. In *Proceedings of the 2008 Int’l Symposium on Information Theory (ISIT)*, pp. 890-894, 2008.

“The Hiring Problem and Lake Wobegon Strategies,” with A. Broder, A. Kirsch, R. Kumar, E. Upfal, and S. Vassilvitskii. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 1184-1193, 2008. Journal version: *SIAM Journal on Computing*, 39(4), pp. 1233-1255, 2009.

“Why Simple Hash Functions Work: Exploiting the Entropy in a Data Stream,” with S. Vadhan. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 746-755, 2008. Journal version: with K. Chung and S. Vadhan. *Theory of Computing*, 9, pp. 897-945-1255, 2013.

“Trace reconstruction with constant deletion probability and related results,” with T. Holenstein, R. Panigrahy, and U. Wieder. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 389-398, 2008.

“Capacity bounds for sticky channels.” *IEEE Transactions on Information Theory*, vol 54, Issue 1, pp. 72-77, 2008.

“Network Monitoring Using Traffic Dispersion Graphs,” with M. Iliofotou, P. Pappu, G. Varghese, M. Faloutsos, and S. Singh. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pp. 315-320, 2007.

“HEXA: Compact Data Structures for Faster Packet Processing,” with S. Kumar, J. Turner, and P. Crowley. In *Proceedings of the IEEE International Conference on Network Protocols*, pp. 246-255, 2007.

“Using a Queue to De-amortize Cuckoo Hashing in Hardware,” with A. Kirsch and M. Mitzenmacher. In *Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing*, 2007.

“Improved lower bounds for the capacity of i.i.d. deletion and duplication channels, with E. Drinea. *IEEE Transactions on Information Theory*, vol 53, Issue 8, pp. 2693-2714, 2007.

“Wired Geometric Routing,” with J. Ledlie, P. Pietzuch, and M. Seltzer. In *Proceedings of the 6th International Workshop on Peer-to-Peer Systems (IPTPS)*, 2007.

“Capacity Upper Bounds for Deletion Channels,” with S. Diggavi and H. Pfister. In *Proceedings of the 2007 Int’l Symposium on Information Theory (ISIT)*, pp. 1716-1720, 2007.

“Codes for Deletion and Insertion Channels with Segmented Errors,” with Z. Liu. In *Proceedings of the 2007 Int’l Symposium on Information Theory (ISIT)*, pp. 846-850, 2007. Journal version: *IEEE Transactions on Information Theory*, vol 56, Issue 1, pp. 224-232, 2010.

“Towards a theory of networked computation : Executive Summary,” with J. Feigenbaum. *ACM SIGACT News*, (4):22-26, 2006.

“Bloom Filters via d-Left Hashing and Dynamic Bit Reassignment,” with F. Bonomi, R. Panigrahy, S. Singh, and G. Varghese. In *Proceedings of the Allerton Conference*, pp. 877-883, 2006.

- “Beyond Bloom Filters: From Approximate Membership Checks to Approximate State Machines,” with F. Bonomi, R. Panigrahy, S. Singh, and G. Varghese. In *Proceedings of ACM SIGCOMM*, pp. 315-326, 2006.
- “An Improved Construction for Counting Bloom Filters,” with F. Bonomi, R. Panigrahy, S. Singh, and G. Varghese. In *Proceedings of the European Symposium on Algorithms*, pp. 684-695, 2006.
- “Less Hashing, Same Performance: Building A Better Bloom Filter,” with A. Kirsch. In *Proceedings of the European Symposium on Algorithms*, pp. 456-467, 2006. Journal version: *Random Structures and Algorithms*, 33(2), pp. 187-218, 2008.
- “Stochastic Shortest Paths via Quasi-Convex Maximization,” with M. Brand, J. Kelner, and E. Nikolova. In *Proceedings of the European Symposium on Algorithms*, pp. 552-563, 2006.
- “Polynomial Time Low-Density Parity-Check Codes with Rates Very Close to the Capacity of the q -ary Random Deletion Channel for Large q .” *IEEE Transactions on Information Theory*, vol 52, Issue 12, pp. 5496-5501, 2006.
- “A Simple Lower Bound for the Capacity of the Deletion Channel,” with E. Drinea. *IEEE Transactions on Information Theory*, vol 52, Issue 10, pp. 4657-4660, 2006.
- “On Lower Bounds for the Capacity of Deletion Channels,” with E. Drinea. *IEEE Transactions on Information Theory*, vol 52, Issue 10, pp. 4648-4657, 2006.
- “On the Theory and Practice of Data Recovery with Multiple Versions.” In *Proceedings of the International Symposium on Information Theory*, pp. 982-986, 2006.
- “Fine-Grained Layered Multicast with STAIR,” with J. Byers, Gu-In Kwon, and M. Luby. *IEEE/ACM Transactions on Networking*, 14:1, pp. 81-93, 2006.
- “Distance-Sensitive Bloom Filters,” with A. Kirsch. In *Proceedings of the ALENEX Conference*, pp. 41-50, 2006.
- “BubbleSearch: A Simple Heuristic for Improving Priority-Based Greedy Algorithms,” with N. Lesh. *Information Processing Letters*, 97(4), pp. 161-169, 2006.
- “Network-Aware Overlays with Network Coordinates,” with P. Pietzuch, J. Ledlie, and M. Seltzer. In *Proceedings of the International Workshop on Dynamic Distributed Systems*, 2006.
- “Editorial: The Future of Power Law Research.” *Internet Mathematics*, vol. 2. no. 4, pp. 525-534, 2006.
- “X-Tolerant Test Response Compaction,” with S. Mitra, S. Lumetta, and N. Patil. *IEEE Design and Test of Computers*, vol. 22, no. 6, pp. 566-574, Nov/Dec 2005.
- “Privacy preserving keyword searches on remote encrypted data,” with Y.-C. Chang. In *Proceedings of the Third Intl. Conference on Applied Cryptography and Network Security*, pp. 442-455, 2005.
- “Simple Summaries for Multiple Choice Hashing,” with A. Kirsch. In *Proceedings of the 43rd Annual Allerton Conference on Communication, Control, and Computing*, 2005. Journal version: *IEEE/ACM Transactions on Networking*, 16(1), pp. 218-231, 2008.

- “A case study in large-scale interactive optimization,” with M. Chimani, N. Lesh, and C. Sidner. In *Proceedings of Artificial Intelligence and Applications 2005*, pp. 24-29.
- “New heuristic and interactive approaches to 2D rectangular strip packing,” with N. Lesh, A. McMahon, and J. Marks. *ACM Journal of Experimental Algorithms*, 10, 2005.
- “Multidimensional Balanced Allocations,” with A. Broder. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 195-196, 2005.
- “Digital Fountains: A Survey and Look Forward.” 2004 Information Theory Workshop.
- “Improved Lower Bounds on the Capacity of I.I.D. Deletion Channels,” with E. Drinea. In *Proceedings of the 42nd Annual Allerton Conference on Communication, Control, and Computing*, 2004.
- “Improved Lower Bounds on the Capacity of Deletion Channels,” with E. Drinea. In *Proceedings of the 2004 IEEE International Symposium on Information Theory*.
- “Geometric Generalizations of the Power of Two Choices,” with J. Byers and J. Conside. In *Proceedings of the 16th ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pp. 54-63, 2004.
- “Interactive Data Summarization: An Example Application,” with N. Lesh. In *Proceedings of the 2004 Workshop on Advanced Visual Interfaces*, pp. 183-187.
- “On Power Law Distributions for Monkeys Typing Randomly: The Case of Unequal Probabilities,” with B. Conrad. *IEEE Transactions on Information Theory*, 50:7, pp. 1403-1414, 2004.
- “Dynamic Models for File Sizes and Double Pareto Distributions.” *Internet Mathematics*, 1(3):305-334, 2003.
- “A Scaling Result for Explosive Processes,” with R. Oliveira and J. Spencer. *Electronic Journal of Combinatorics*, 11(1), R31, 2004.
- “An Exhaustive Approach to 2D Strip Packing,” with N. Lesh, A. McMahon, and J. Marks. *Information Processing Letters*, 90(1), pp. 7-14, 2004.
- “New Models and Methods for File Size Distributions,” with B. Twestzky. In *Proceedings of the 41st Annual Allerton Conference on Communication, Control, and Computing*, pp. 603-612, 2003.
- “Capacity Approaching Signal Constellations for Channels with Memory,” with A. Kavčić, and X. Ma. *IEEE Transactions on Information Theory*, 49:7, pp. 1636-1652, 2003.
- “The Power Spectra of Good Codes for Partial Response Channels,” with A. Kavcic and X. Ma. In *Proceedings of the 2003 IEEE International Symposium on Information Theory*, p. 45.
- “Concatenated Codes for Deletion Channels,” with J. Chen, C. Ng, and N. Varnica. In *Proceedings of the 2003 IEEE International Symposium on Information Theory*, p. 218.
- “Verification Codes for Deletion Channels.” In *Proceedings of the 2003 IEEE International Symposium on Information Theory*, p. 217.

- “Estimating and Comparing Entropies Across Written Natural Languages Using PPM Compression,” with F. Behr, V. Fossum, and D. Xiao. In *Proceedings of the 2003 IEEE Data Compression Conference*, p. 416.
- “Simple Load Balancing for Distributed Hash Tables,” with J. Byers, and J. Considine. Appears in the 2nd International Workshop on Peer-to-Peer Systems (IPTPS), pp. 80-87, 2003.
- “A Complete and Effective Move Set for Simplified Protein Folding,” with N. Lesh, and S. Whitesides. In *Proceedings of the 7th Annual International Conference on Research in Computational Molecular Biology*, pp 188-195, 2003.
- “Network Applications of Bloom Filters: A Survey,” with A. Broder. In *Proceedings of the 40th Annual Allerton Conference on Communication, Control, and Computing*, pp. 636-646, 2002. Journal version: *Internet Mathematics*, 1(4):485-509, 2003.
- “Verification Codes,” with M. Luby. In *Proceedings of the 40th Annual Allerton Conference on Communication, Control, and Computing*, pp. 38-47, 2002. Journal version: *IEEE Transactions on Information Theory*: 51(1), pp. 120-127, 2005.
- “Balls and Bins with Memory,” with B. Prabhakar and D. Shah. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 799-808, 2002.
- “Informed Content Delivery Across Adaptive Overlay Networks,” with J. Byers, J. Considine, and S. Rost. In *Proceedings of ACM SIGCOMM '02*, available as *Computer Communications Review*, vol. 32:4, pp. 47-60, 2002. Journal version: *IEEE/ACM Transactions on Networking*, 12:5, pp. 767-780, 2004.
- “Optimal Plans for Aggregation,” with A. Broder. In *Proceedings of the 21st Annual ACM Symposium on Principles of Distributed Computing*, pp. 144-152, 2002.
- “Human-Guided Tabu Search,” with G. Klau, N. Lesh, and J. Marks. In *Proceedings of the 18th National Conference on Artificial Intelligence*, pp. 41-47, 2002.
- “The HuGS Platform: A Toolkit for Interactive Optimization,” with G. Klau, N. Lesh, J. Marks, and G. Schafer. In *Proceedings of the Workshop on Advanced Visual Interfaces*, pp. 324-330, 2002.
- “Exact Sampling of TCP Window States,” with A. Goel. In *Proceedings of IEEE INFOCOM*, pp. 259-265, 2002.
- “Balls and Bins Models with Feedback,” with E. Drinea and A. Frieze. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 308-315, 2002.
- “A Brief History of Lognormal and Power Law Distributions.” In *Proceedings of the 39th Annual Allerton Conference on Communication, Control, and Computing*, pp. 182-191, 2001. Journal version: *Internet Mathematics*, 1(2):226-251, 2003.
- “Capacity Approaching Signal Constellations for Channels with Memory,” with A. Kavčić, X. Ma, and N. Varnica. In *Proceedings of the 39th Annual Allerton Conference on Communication, Control, and Computing*, pp. 311-320, 2001.
- “Compressed Bloom Filters.” In *Proceedings of the 20th Annual ACM Symposium on Principles of Distributed Computing*, pp. 144-150, 2001. Journal version: *IEEE/ACM Transactions on Networks*, 10:5, pp. 613-620, October 2002.

“Towards More Completed Models of TCP Throughput,” with R. Rajaraman. *Journal of Supercomputing*, 20(2), pp. 137-160, September 2001.

“Deriving Performance Bounds for ISI Channels Using Gallager Codes,” with A. Kavčić, B. Marcus, and B. Wilson. In *Proceedings of the 2001 International Symposium on Information Theory*, p. 345, 2001.

“Toward Compressing Web Graphs,” with M. Adler. In *Proceedings of the 2001 Data Compression Conference*, pp. 203-212, 2001.

“On the Hardness of Finding Multiple Preset Dictionaries.” In *Proceedings of the 2001 Data Compression Conference*, pp. 411-418, 2001. Journal version: *IEEE Transactions on Information Theory*, 50:7, pp. 1536-1539, 2004.

“Fine-Grained Layered Multicast,” with J. Byers and M. Luby. In *Proceedings of IEEE INFOCOM 2001*, pp. 1143-1151, 2001. Journal version: “Fine-Grained Layered Multicast with STAIR,” with J. Byers, Gu-In Kwon, and M. Luby. To appear in *IEEE/ACM Transactions on Networking*.

“Using Multiple Hash Functions to Improve IP Lookups,” with A. Broder. In *Proceedings of IEEE INFOCOM 2001*, pp. 1454-1463, 2001.

“Improved Results for Route Planning in Stochastic Transportation Networks,” with J. Boyan. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 895-902, 2001.

“Estimating Resemblance of MIDI Documents,” with S. Owen. In *Proceedings of the 3rd Workshop on Algorithm Engineering and Experiments*, available as *Lecture Notes in Computer Science 2153*, pp. 78-90, 2001.

“Linear Waste of Best Fit Bin Packing on Skewed Distributions,” with Claire Kenyon. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pp. 582-589, 2000. Journal version: invited to special issue of *Random Structures and Algorithms*, 20(3), pp. 441-464, 2002.

“FLID/DL: Congestion Control for Layered Multicast,” with J. Byers, M. Frumin, G. Horn, M. Luby, A. Roetter, and W. Shaver. In *Proceedings of the 2nd International Workshop on Networked Group Communication (NGC 2000)*, pp. 71-81, 2000. Journal version: with J. Byers, G. Horn, M. Luby, and W. Shaver. *IEEE Journal on Selected Areas in Communications*, 20:8, pp. 1558-1570, October 2002.

“On Near-Uniform URL Sampling,” with M. Henzinger, A. Heydon, and M. Najork. In *Proceedings of the Ninth International World Wide Web Conference*, pp. 295-308, 2000.

“An Extension of Path Coupling and Its Application to the Glauber Dynamics for Path Coupling,” with M. Dyer, L. Goldberg, C. Greenhill, and M. Jerrum. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 616-623, 2000. Journal version: *SIAM Journal on Computing*, 30(6), pp. 1962-1975, 2001.

“Improved Classification via Connectivity Information,” with A. Broder and R. Krauthgamer. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 576-585, 2000.

“The Asymptotics of Selecting the Shortest of Two, Improved,” with B. Vöcking. In *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, pp. 326-327, 1999. Extended version appears as a chapter in *Analytic Methods in Applied Probability: In Memory of Fridrikh Karpelevich*, American Mathematical Society, pp. 165-176, 2002.

“Completeness and Robustness Properties of Min-Wise Independent Permutations,” with A. Broder. In *Proceedings of Random '99*, available as *Lecture Notes in Computer Science 1671*, pp. 1-10, 1999. Journal version: *Random Structures and Algorithms*, 18(1), pp. 18-30, January 2001.

“Measuring Index Quality Using Random Walks on the Web,” with M. Henzinger, A. Heydon, and M. Najork. In *Proceedings of the Eighth International World Wide Web Conference*, pp. 213-225, Elsevier Science, 1999.

“Analysis of Timing Based Mutual Exclusion with Random Times,” with E. Gafni. In *Proceedings of the 18th Annual ACM Symposium on Principles of Distributed Computing*, pp. 13-21, 1999. Journal version: *SIAM Journal on Computing*, 31:3, pp. 816-837, 2001.

“Accessing Multiple Mirror Sites in Parallel: Using Tornado Codes to Speed Up Downloads,” with J. Byers and M. Luby. In *Proceedings of the 18th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '99)*, pp. 275-284, 1999.

“Unscrambling Address Lines (Abstract),” with A. Broder and L. Moll. In *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 870-871, 1999.

“A Derandomization Using Min-Wise Independent Permutations,” with A. Broder and M. Charikar. In *Proceedings of Random '98*, available as *Lecture Notes in Computer Science 1518*, pp. 15-24, October 1998. Journal version: *Journal of Discrete Algorithms*, 1:1, pp. 11-20, 2003.

“On Balls and Bins with Deletions,” with R. Cole, A. Frieze, B. M. Maggs, A. W. Richa, R. Sitaraman, and E. Upfal. In *Proceedings of Random '98*, available as *Lecture Notes in Computer Science 1518*, pp. 145-158, October 1998.

“Studying Balanced Allocations with Differential Equations.” *Combinatorics, Probability, and Computing*, vol. 8, pp. 473-482, 1999.

“Delayed Information and Action in On-Line Algorithms,” with S. Albers and M. Charikar. In *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science (FOCS-98)*, pp. 71-80, November 1998. Journal version: *Information and Computation*, 170, pp. 135-152, 2001.

“A Digital Fountain Approach to the Reliable Distribution of Bulk Data,” with John Byers, Michael Luby, and Ashu Rege. In *Proceedings of ACM SIGCOMM '98*, available as *Computer Communications Review*, vol. 28:4, pp. 56-67, 1998. Journal version: *IEEE Journal on Selected Areas in Communication*, 20:8, pp. 1528-1540, October 2002.

“Analyses of Load Stealing Models Using Differential Equations.” In *Proceedings of the 10th Annual ACM Symposium on Parallel Algorithms and Architectures*, pp. 212-221, 1998. Journal version: *Theory of Computing Systems*, 34:1, pp. 77-98, January 2001.

“Load Adaptive Routing on Butterfly Networks,” with R. Cole, B. Maggs, F. Meyer auf der Heide, A. Richa, K. Schröder, R. Sitaraman, and B. Vöcking. In *Proceedings of the*

30th Annual ACM Symposium on Theory of Computing (STOC-98), pp. 378-388, ACM Press, May 23-26, 1998.

“Min-Wise Independent Permutations,” with Andrei Broder, Moses Charikar, and Alan Frieze. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pp. 327-336, ACM Press, May 23-26, 1998. Journal version: *Journal of Computer and System Sciences*, 60(3), pp. 630-659, 2000.

“Analysis of Low Density Codes and Improved Designs Using Irregular Graphs,” with Michael Luby, Amin Shokrollahi, and Dan Spielman. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pp. 249-258, ACM Press, May 23-26, 1998.

“Improved Low Density Parity Check Codes Using Irregular Graphs and Belief Propagation,” with Michael Luby, Amin Shokrollahi, and Dan Spielman. In *Proceedings of the 1998 International Symposium on Information Theory*. Extended version available as Compaq Systems Research Center Technical Note 1998-009. Journal version: “Improved Low Density Parity Check Codes Using Irregular Graphs,” *IEEE Transactions on Information Theory*, 47(2), pp. 585-598, 2001.

“Analysis of Random Processes via And-Or Tree Evaluation,” with Michael Luby and Amin Shokrollahi. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 364-373, 1998.

“Average Case Analyses of First Fit and Random Fit Bin Packing,” with Susanne Albers. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 290-299, 1998. Journal version: *Random Structures and Algorithms*, 16(3), pp. 240-259, 2000.

“How Useful is Old Information?” In *Proceedings of the 16th ACM Symposium on Principles of Distributed Computing*, pp. 83-91, 1997. Journal version: *IEEE Transactions on Parallel and Distributed Systems*, 11:1, pp. 6-20, January 2000.

“Revisiting the COUNTER Algorithms for List Update,” with Susanne Albers. *Inf. Proc. Letters*, 64, pp. 155-160, 1997.

“Practical Loss-Resilient Codes,” with Michael Luby, Amin Shokrollahi, Dan Spielman, and Volker Stemann. In *Proceedings of the 29th ACM Symposium on the Theory of Computing*, pp. 150-159, 1997. Journal version: “Efficient Erasure Correcting Codes,” with Michael Luby, Amin Shokrollahi, and Dan Spielman. *IEEE Transactions on Information Theory*, 47(2), pp 569-584, 2001.

“On the Analysis of Randomized Load Balancing Schemes.” In *Proceedings of the 9th ACM Symposium on Parallel Algorithms and Architectures*, pp. 292-301, 1997. Journal version: special conference issue of *Theory of Computing Systems*, vol. 32, pp. 361-386, 1999.

“Load Balancing and Density Dependent Jump Markov Processes.” In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, pp. 213-222, 1996. Journal version: “The Power of Two Choices in Randomized Load Balancing.” *IEEE Transactions on Parallel and Distributed Computing*, 12(10), pp. 1094-1104, 2001.

“Designing Stimulating Programming Assignments for an Algorithms Course: A Collection of Problems Based on Random Graphs.” In *SIGCSE Bulletin*, 28:3, pp. 29-36, September 1996.

“Constant Time per Edge is Optimal in Rooted Tree Networks.” In *Proceedings of the 8th ACM Symposium on Parallel Algorithms and Architectures*, pp. 162-169, 1996. Journal version: *Distributed Computing* vol. 10:4, 1997, pp. 189-197.

“Average Case Analyses of List Update Algorithms, with Applications to Data Compression,” with Susanne Albers. In *Proceedings of the 1996 International Colloquium on Automata, Languages, and Programming*, pp. 514-525. Journal version: *Algorithmica*, vol. 21:3, 1998, pp. 312-329.

“Pattern-based Compression of Text Images,” with Andrei Broder. In *1996 Data Compression Conference*, pp. 300-309.

“Parallel Randomized Load Balancing,” with Micah Adler, Soumen Chakrabarti, and Lars Rasmussen. In *Proceedings of the 27th ACM Symposium on the Theory of Computing*, pp. 238-247, 1995. Journal version: *Random Structures and Algorithms*, vol. 13:2, September 1998, pp. 159-188.

“Bounds on the Greedy Routing Algorithm in Array Networks.” In *Proceedings of the 6th ACM Symposium on Parallel Algorithms and Architectures*, pp. 346-353, 1994. Journal version: special conference issue of the *Journal of Computer and System Sciences* vol. 53:3, December 1996, pp. 317-327.

“Computational Complexity of Loss Networks,” with Graham Louth and Frank Kelly. *Theoretical Computer Science*, vol. 125, May 1994, pp. 45-59.

Other Publications

“Lyra: An Efficient and Expressive Subquadratic Architecture for Modeling Biological Sequences,” with Krithik Ramesh, Sameed Siddiqui, Albert Gu, and Pardis Sabeti. Available at <https://arxiv.org/abs/2503.16351>.

“Juniper: Reconstructing Transmission Events from Next-Generation Sequencing Data at Scale,” with Ivan Specht, Gage Moreno, Taylor Brock-Fisher, Lydia Kraslinkova, Brittany Petros, Jonathan Pekar, Mark Schifferli, Ben Fry, Catherine Brown, Lawrence Madoff, Megan Burns, Stephen Schaffner, Daniel Park, Bronwn MacInnis, Al Ozonoff, Patrick Varilly, and Pardis Sabeti. Available at <https://www.medrxiv.org/content/10.1101/2025.03.02.25323192v1>.

“Delphy: scalable, near-real-time Bayesian phylogenetics for outbreaks,” with Patrik Varilly, Mark Schifferli, Katherine Yang, Tim Burcham, Paul Cronan, Olivia Glennon, Olivia Jacks, Ellory Laning, Libby Marrs, Kyle Oba, Shannon Yeung, Edyth Parker, Ifeanyi Omah, Jonathan E. Pekar, Laura Luebbert, Kristian G. Andersen, Daniel J. Park, Stephen F. Schaffner, Bronwyn L. MacInnis, Christian Happi, Jacob E. Lemieux, Al Ozonoff, Michael D. Mitzenmacher, Ben Fry, and Pardis C. Sabeti. Available at <https://www.biorxiv.org/content/10.1101/2025.03.25.645253v1>.

“Use of popularity information to reduce risk posed by guessing attacks,” with Stuart Schechter and Cormac Herley. US Patent 08959644. Issued 2/17/2015.

“Method and apparatus providing network coding based flow control,” with Jay Kumar Sundararajan, Devavrat Shah, Muriel Medard, Joao Barros, and Szymon Jakubczak. US Patent 08526451. Issued 9/3/2013.

“Method and apparatus providing network coding based flow control,” with Jay Kumar Sundararajan, Devavrat Shah, Muriel Medard, Joao Barros, and Szymon Jakubczak. US Patent 08130776. Issued 3/6/2012.

“Method for finding optimal paths using a stochastic network model,” with Evdokia Nikolova and Matthew Brand. US Patent 07573866. Issued 8/11/2009.

“Ranking search engine results,” with Monika Henzinger. US Patent 07541388. Issued 11/11/2008.

“Distributed, compressed Bloom filter Web cache server.” US Patent 06920477. Issued 7/19/2005.

“Method for packing rectangular strips,” with Neal Lesh and Joe Marks. US Patent 06832129. Issued 12/14/2004.

“System and method for near-uniform sampling of web page addresses,” with Monika Henzinger, Allan Heydon, and Marc Najork. US Patent 06594694. Issued 7/15/2003.

“On demand encoding with a window,” with Michael Luby, Gavin Horn, Jeffrey Persch, John Byers, and Armin Haken. US Patent 06486803. Issued 11/26/2002.

“Generating high weight encoding symbols using a basis,” with Armin Haken, Michael Luby, Gavin Horn, Diane Hernek, and John Byers. US Patent 06411223. Issued 6/25/2002.

“Space-efficient multi-cycle barrel shifter circuit,” with Laurent Moll. US Patent 06314156. Issued 11/6/2001.

“Method for determining a random permutation of variables by applying a test function,” with Andrei Broder, Laurent Moll, and Mark Shand. US Patent 06292762. Issued 9/18/2001.

“Loss resilient code with double heavy tailed series of redundant layers,” with Michael Luby. US Patent 06195777. Issued 2/27/2001.

“Message encoding with irregular graphing,” with Michael Luby, Amin Shokrollahi, Dan Spielman, and Volker Stemann. US Patent 06163870. Issued 12/19/2000.

“Compression of grey scale images of text,” with Andrei Broder. US Patent 06088309. Issued 7/11/2000.

“Irregularly graphed encoding technique,” with Michael Luby, Amin Shokrollahi, Dan Spielman, and Volker Stemann. US Patent 06081909. Issued 6/27/2000.

“Loss resilient decoding technique,” with Michael Luby, Amin Shokrollahi, Dan Spielman, and Volker Stemann. US Patent 06073250. Issued 6/6/2000.

“Task processing optimization in a multiprocessor system,” with Andrei Broder. US Patent 05991808. Issued 11/23/1999.

“Compression protocol with multiple preset dictionaries,” with Andrei Broder and Jeff Mogul. US Patent 05953503. Issued 9/14/1999.

“When Simple Hash Functions Suffice,” with K. Chung and S. Vadhan. Appears as a chapter in *Beyond Worst Case Analysis*, edited by Tim Roughgarden, Cambridge University Press.

“Algorithms with Predictions,” with S. Vassilvitskii. Appears as a chapter in *Beyond Worst Case Analysis*, edited by Tim Roughgarden, Cambridge University Press.

“Some Practical Randomized Algorithms and Data Structures,” with E. Upfal. Appears as a chapter in *Computing Handbook, Third Edition: Computer Science and Software Engineering*, edited by T. Gonzalez, J. Diaz-Herrera, and A. Tucker. Chapman and Hall, 2014.

“Building an Efficient Hash Table on the GPU,” with D. Alcantara, V. Volkov, S. Sengupta, J. Owens, and N. Amenta. Appears as a chapter in *GPU Computing Gems: Jade Edition*, edited by Wen-mei W. Hwu. Morgan Kaufmann, 2011.

“Codes – Protecting Data Against Loss and Errors.” Appears as a chapter in *Algorithms Unplugged*, edited by B. Vöcking, M. Dietzfelbinger, R. Reischuk, and H. Vollmer. Springer, 2010.

“Hash-Based Techniques for High-Speed Packet Processing”, with A. Kirsch and G. Varghese. Appears as a chapter in *Algorithms for Next Generation Networks*, edited by G. Cormode and M. Thottan. Springer, 2010.

“Bloom Filters.” Encyclopedia of Database Systems, 2009.

“Unbiasing Random Bits.” *Dr. Dobbs Journal*, No. 335, pp. 101-104, April 2002.

“Good Hash Tables & Multiple Hash Functions.” *Dr. Dobbs Journal*, No. 336, May 2002.

“Challenging Students with Creative Assignments.” *SIGACT News*, 32(1):70-73, 2001.

“An Experimental Assignment on Random Processes.” *SIGACT News*, 32(1):74-78, 2001.

“The Power of Two Random Choices: A Survey of Techniques and Results,” with A. Richa and R. Sitaraman. Appears as a chapter in *Handbook of Randomized Computing: Volume 1*, edited by P. Pardalos, S. Rajasekaran, and J. Rolim, pp. 255-312.

“Tornado Codes for Forward Error Correction,” with Michael Luby, in *Digital Forefront*.

“Bounds and Improvements for BiBa Signature Schemes,” with A. Perrig. Harvard Computer Science Group Technical Report TR-02-02.

“A Note on Low Density Parity Check Codes for Erasures and Errors.” Digital SRC Technical Note 1998-017.

“Tight Thresholds for the Pure Literal Rule.” Digital SRC Technical Note 1997-011.

“The Power of Two Choices in Randomized Load Balancing.” Ph.D. thesis, U.C. Berkeley, August 1996.

“*Calculus: A Solution Manual*,” with others, Harcourt Brace Jovanovich, 1993.

“*Elliptic Curves in Computer Science: Primality Testing, Factoring, and Cryptography*,” senior honors thesis, Harvard University, 1991. Published as a Hoopes Prize winner.

Conference Presentations “Queueing, Predictions, and LLMs: Challenges and Open Problems.” Tutorial, SIGMETRICS. June 2025.

“SkipPredict: When to Invest in Predictions for Scheduling. Presented at the TTIC Workshop on Learning-Augmented Algorithms, August 2024.

“Algorithms with Predictions.” Invited talk presented at the INFORMS 2023 Annual Meeting, October 2023.

“Algorithms with Predictions.” Presented at the 2nd Vienna Workshop on Computational Optimization. March 2023.

“Algorithms with Predictions.” Presented at the Sublinear Algorithms Workshop, FODSI/MIT. August 2022.

“Algorithms with Predictions: State of the Area.” Presented at the Algorithms with Predictions Workshop at the 54th Annual ACM Symposium on Theory of Computing. June 2022. Also panel lead.

“Efficient Randomized Network Telemetry Algorithms.” Presented at the Data plane algorithms in programmable networks tutorial at ACM SIGMETRICS 2022.

“Scheduling with Predictions and the Price of Misprediction.” Presented at the 11th Annual Innovations in Theoretical Computer Science Conference. January 2020.

“Learning-Augmented Algorithms: How ML Can Lead to Provably Better Algorithms.” ALGO Keynote Talk, September 2019.

“Robust Set Reconciliation via Locality Sensitive Hashing.” Presented at the 38th Symposium on Principles of Database Systems, July 2019.

“Arithmetic Progression Hypergraphs: Examining the Second Moment Method.” Presented at the 2019 ANALCO Workshop, January 2019.

“A Model for Learned Bloom Filters and Optimizing by Sandwiching.” Presented at the NeurIPS 2018 Poster Session.

“Adaptive Cuckoo Filters.” Presented at the 2018 ALENEX Conference, January 2018.

“Analyzing Distributed Join-Idle-Queue: A Fluid Limit Approach.” Presented at the 54th Annual Allerton Conference on Communication, Control, and Computing, October 2016.

“Voronoi Choice Games.” Presented at the 2016 International Colloquium on Automata, Languages, and Programming, July 2016.

“A New Approach to Analyzing Robin Hood Hashing.” Presented at the 2016 ANALCO Workshop, January 2016.

“More Analysis of Double Hashing for Balanced Allocations.” Presented at the 2016 ANALCO Workshop, January 2016.

“Data exploration: equitability and the maximal information coefficient.” Presented at the Workshop for Data-driven Algorithmics, September 2015.

“Multi-Party Set Reconciliation Using Characteristic Polynomials.” Presented at the 2014 Allerton Conference, October 2014.

“Parallel Peeling Algorithms.” Presented at the 2014 Symposium on Parallelism in Algorithms and Architectures, June 2014.

“Balanced Allocations and Double Hashing.” Presented at the 2014 Symposium on Parallelism in Algorithms and Architectures, June 2014.

“Peeling Arguments and Double Hashing.” Presented at the 2012 Allerton Conference, October 2012.

“The Complexity of Object Reconciliation and Open Problems Related to Set Difference and Coding.” Presented at the 2012 Allerton conference, October 2012.

“The Ecology of Communication, Control, and Computing Research and Education (Panel).” Presented at the 2012 Allerton conference, October 2012.

“Peeling Arguments: Invertible Bloom Lookup Tables and Biff Codes.” Presented at the 2012 Workshop on Algorithms for Modern Massive Data Sets (MMDS), July 2012.

“Invertible Bloom Lookup Tables.” Presented at the 2011 Allerton conference, September 2011.

“Brief Announcement: Large-Scale Multimaps.” Presented at the 2011 Symposium on Parallel Algorithms and Architectures, June 2011.

“On the Zero-Error Capacity Threshold for Deletion Channels.” Presented at the 2011 ITA Workshop, February 2011.

“Some Recent Uses of Hashing for Network Applications.” Presented at the 2010 Workshop on Research and Analysis of Tail Phenomena, August 2010.

“Some Recent Uses of Hashing for Network Applications.” Presented at the 2010 Workshop on Random Graphs and Algorithms, June 2010.

“Tight Asymptotic Bounds for the Deletion Channel with Small Deletion Probabilities.” Presented at the 2010 International Symposium on Information Theory, June 2010.

“Some Open Questions on Cuckoo Hashing.” Presented at the 2009 European Symposium on Algorithms, August 2009.

“On the Performance of Multiple Choice Hash Tables with Moves on Deletions and Inserts.” Presented at the 2008 Allerton conference, September 2008.

“Designing Floating Codes for Expected Performance.” Presented at the 2008 Allerton conference, September 2008.

“A Survey of Results for Deletion Channels and Related Synchronization Channels.” Presented at the 11th Scandinavian Workshop on Algorithm Theory, July 2008.

“Why Simple Hash Functions Work: Exploiting the Entropy in a Data Stream.” Presented at the 19th Annual ACM-SIAM Symposium on Discrete Algorithms, January 2008.

“Using a Queue to De-amortize Cuckoo Hashing in Hardware.” Presented at the 2007 Allerton conference, September 2007.

“Capacity Upper Bounds for Deletion Channels.” Presented at the 2007 Int’l Symposium on Information Theory (ISIT), June 2007.

“Codes for Deletion and Insertion Channels with Segmented Errors.” Presented at the 2007 Int’l Symposium on Information Theory (ISIT), June 2007.

“Digital Fountains: Applications and Related Issues.” Presented at the INFORMS Annual Meeting, November 2006.

“New Directions for Power Law Research.” Presented at the Radcliffe Exploratory Seminar on Dynamic Networks, October 2006.

“Bloom Filters via d -left Hashing and Dynamic Bit Reassignment.” Presented at the 2006 Allerton conference, September 2006.

“Beyond Bloom Filters: From Approximate Membership Checks to Approximate State Machines.” Presented at the 2006 SIGCOMM conference, September 2006.

“An Improved Construction for Counting Bloom Filters.” Presented at the European Symposium on Algorithms, September 2006.

“Building a Better Bloom Filter.” Presented at the European Symposium on Algorithms, September 2006.

“On the Theory and Practice of Data Recovery with Multiple Versions.” Presented at the 2006 International Symposium on Information Theory, July 2006.

“New Results for an Old Data Structure: Bloom Filters and Beyond.” Presented at the 2006 Stochastic Networks conference, June 2006.

“New Results and Open Problems for Deletion Channels.” Presented at ANALCO 2006, January 2006 (plenary talk).

“Digital Fountains and their Application to Informed Content Delivery over Adaptive Overlay Networks.” Presented at 19th International Symposium on Distributed Computing, September 2005 (plenary talk).

“Simple Summaries for Multiple Choice Hashing.” Presented at the 2005 Allerton Conference.

“New Directions in Power Law Research.” MSRI Workshop on Models of Real-World Random Networks, April 2005.

“Multi-dimensional Balanced Allocations.” Presented at the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, January 2005.

“Digital Fountains: A Survey and Look Forward.” Presented at the 2004 Information Theory Workshop. October 2004.

“Improved Lower Bounds on the Capacity of I.I.D. Deletion Channels.” Presented at the 2004 Allerton Conference.

“Geometric Generalizations of the Power of Two Choices.” Presented at the 16th ACM Symposium on Parallel Algorithms and Architectures, June 2004.

“New Models and Methods for File Size Distributions.” Presented at the 40th Annual Allerton Conference on Communication, Control, and Computing, October 2003.

“Verification Codes for Deletion Channels.” Presented at the International Symposium on Information Theory, July 2003.

“New Exhaustive, Heuristic, and Interactive Algorithms for 2-Dimensional Bin-Packing.” Presented at the 2nd Annual Workshop on Random Graphs and Algorithms, June 2003.

“Network Applications of Bloom Filters: A Survey.” Presented at the 40th Annual Allerton Conference on Communication, Control, and Computing, October 2002.

“Verification Codes.” Presented at the 40th Annual Allerton Conference on Communication, Control, and Computing, October 2002.

“New Directions in Balls and Bins Problems.” Invited talk for the 6th International Workshop on Randomization and Approximation Methods in Computer Science. September 2002.

“Optimal Plans for Aggregation.” Presented at the 21st Annual ACM Symposium on Principles of Distributed Computing, July 2002.

“How to Get Across Town by Bus.” Presented at the INFORMS Annual Meeting, November 2001.

“A Brief History of Lognormal and Power Law Distributions.” Presented at the 39th Annual Allerton Conference on Communication, Control, and Computing, October 2001.

“Compressed Bloom Filters.” Presented at the 20th Annual ACM Symposium on Principles of Distributed Computing, August 2001.

“How Useful Is Old Information?” Presented at the INFORMS Applied Probability Conference, July 2001.

“Using Multiple Hash Functions to Improve IP Lookups.” Presented at the INFOCOM, April 2001.

“On the Hardness of Finding Multiple Preset Dictionaries.” Presented at the 2001 Data Compression Conference, March 2001.

“Improved Results for Route Planning in Stochastic Transportation Networks.” Presented at the 12th Annual ACM-SIAM Symposium on Discrete Algorithms, January 2001.

“Estimating Resemblance of MIDI Documents.” Presented at the 3rd Workshop on Algorithm Engineering and Experiments, January 2001.

“Linear Waste of Best Fit Bin Packing on Skewed Distributions.” Presented at the 41st Annual IEEE Symposium on Foundations of Computer Science, November 2000.

“Improved Classification via Connectivity Information.” Presented at the 11th Annual ACM Symposium on Discrete Algorithms, January 2000.

“The Asymptotics of Selecting the Shortest of Two, Improved.” Presented at the 37th Annual Allerton Conference on Communication, Control, and Computing, August 1999.

“Analysis of Timing Based Mutual Exclusion with Random Times.” Presented at the 18th Annual ACM Symposium on Principles of Distributed Computing, May 1999.

“Accessing Multiple Mirror Sites in Parallel: Using Tornado Codes to Speed Up Downloads.” Presented at INFOCOM '99, March 1999.

“Delayed Information and Action in On-Line Algorithms.” Presented at the 39th Annual IEEE Symposium on Foundations of Computer Science, November 1998.

“A Derandomization Using Min-Wise Independent Permutations.” Presented at Random ’98, October 1998.

“On Balls and Bins with Deletions.” Presented at Random ’98, October 1998.

“The Power of Two Choices and Other Examples of Using Differential Equations to Analyze Algorithms.” Invited presentation at the Warwick Randomized Algorithms and Stochastic Simulation Tutorial and Workshop. July 1988.

“Analysis of Low Density Codes and Improved Designs Using Irregular Graphs.” Presented at the 30th Annual ACM Symposium on Theory of Computing, May 1998.

“Average Case Analyses of First Fit and Random Fit Bin Packing.” Presented at the 9th Annual ACM Symposium on Discrete Algorithms, January 1998.

“How Useful Is Old Information?” Presented at the 16th Annual ACM Symposium on Principle of Distributed Computing, August 1997.

“On the Analysis of Randomized Load Balancing Schemes.” Presented at the 9th ACM Symposium on Parallel Algorithms and Architectures, Newport, Rhode Island, June 1997.

“Load Balancing and Density Dependent Jump Markov Processes.” Presented at the 37th Annual Symposium on Foundations of Computer Science, Burlington, Vermont, October 1996.

“Constant Time per Edge is Optimal in Rooted Tree Networks.” Presented at the 8th ACM Symposium on Parallel Algorithms and Architectures, Padua, Italy, June 1996.

“GRID: Pattern-based Compression of Text Images.” Presented at the 1996 Data Compression Conference, Snowbird, Utah, April 1996.

“Parallel Randomized Load Balancing.” Presented at the 27th ACM Symposium on the Theory of Computing, Las Vegas, Nevada, May 1995. Extended version presented at the Bay Area Theory Seminar, July 1995.

“Bounds on the Greedy Routing Algorithm on Array Networks.” Presented at the 6th ACM Symposium on Parallel Algorithms and Architectures, Cape May, New Jersey, June 1994.

Other Presentations

“Algorithms with Predictions.” Georgia Tech. 2023.

“Modern Algorithms Lecture Series.” Lipari Summer School. 2022.

“Algorithms with Predictions.” Carnegie Mellon, UC Santa Barbara, Technical University of Munich. 2022.

“Locality Sensitive Hashing.” Broad Institute (Models, Inference, and Algorithms).

“Algorithms with Predictions.” Open Data Science Conference East 2020, Open Data Science Conference Europe 2020, Foundations of Data Science Institute Kickoff Workshop 2020.

“Learning-Augmented Algorithms: How ML Can Lead to Provably Better Algorithms.” Rice University. 2019.

“Metric Sublinear Algorithms via Linear Sampling.” Sublinear Algorithms Workshop (MIT), Baidu Research. 2018.

“Bloom Filters, Cuckoo Hashing, Cuckoo Filters, and Adaptive Cuckoo Filters.” Northeastern University, Univ. of Illinois at Urbana-Champaign, Georgetown University, University of California Irvine, Baidu Research, Georgia Tech. 2018.

“Hashing Lecture Series.” Swedish Summer School in Computer Science. 2016.

“More Analysis of Double Hashing for Balanced Allocations.” U.C. Irvine. 2016.

“Algorithm Engineering with Hashing,” U.C. Irvine, AMD. 2016.

“Bloom Filters and Such,” EADS Summer School on Hashing: Theory and Applications, July 2014.

“Cuckoo Hashing, Balanced Allocations, and Such,” EADS Summer School on Hashing: Theory and Applications, July 2014.

“Peeling Arguments: Invertible Bloom Lookup Tables and Biff Codes,” ICERM Workshop on Stochastic Graph Models, March 2014.

“Peeling Arguments: Invertible Bloom Lookup Tables and Biff Codes,” University of Copenhagen. 2013.

“Some Hash-Based Data Structures and Algorithms Everyone Should Know.” UC Davis, 2013.

“Peeling Algorithms,” Simons Institute. 2013.

“Verifying Computations in the Cloud (and Elsewhere),” TTI/Vanguard : Ginormous Systems. Microsoft Faculty Summit. 2013.

“The Groupon Effect on Yelp Ratings: a Root Cause Analysis,” Cornell, 2012.

“Peeling Arguments: Invertible Bloom Lookup Tables and Biff Codes,” Google. 2012.

“Detecting Novel Associations in Large Data Sets.” eHarmony (LA Machine Learning). 2012.

“Cuckoo Hashing: New Results and Open Questions.” University of Madison-Wisconsin. 2011.

“Practical Verified Computation with Streaming Interactive Proofs.” University of Madison-Wisconsin. 2011.

“The Power of Choice: Simple Hash-Based Data Structures Everyone Should Know.” Santa Fe Institute. 2011.

“Information Asymmetries in Pay-Per-Bid Auctions: How Swoopo Makes Bank.” Liverpool University, Microsoft UK. 2010.

“Some Open Questions on Cuckoo Hashing.” Stanford, Microsoft Research, Yale, Dartmouth, Princeton, Cambridge University, University College London . 2009-2010.

“An Efficient, Rigorous Approach for Identifying Statistically Significant Frequent Itemsets.” Google. 2009.

“Some Recent Results on Cuckoo Hashing.” AT&T. 2009.

“Some Results on Coding for Flash Memory.” University College London. 2009.

“Network Coding Meets TCP.” University College London. 2009.

“A Survey of Results for Deletion Channels and Related Synchronization Channels.” Cambridge, Georgia Tech, Microsoft Research, UCLA, 2009-2010.

“Bloom Filters, Related Data Structures, and their Applications.” Cambridge (Kuwait Lecture). 2009.

“A History of and New Directions for Power Law Research.” SUNY Buffalo. 2009.

“Why Simple Hash Functions Work.” Cisco. 2008.

“Cuckoo Hashing and CAMs.” Cisco, Google. 2008.

“New Results in Hash-Based Data Structures.” Cisco. 2007.

“The Hiring Problem and Lake Wobegon Strategies.” Microsoft, U.C. Berkeley, Princeton, Google, Yahoo. 2007-2008.

“A Brief History of Lognormal and Power Law Distributions.” Harvard Statistics. 2006.

“New Results and Open Problems for Deletion Channels and Related Synchronization Problems.” CMU, Brown, Harvard. 2006.

“Some New Results on an Old Data Structure, Bloom Filters .” Stanford, Cisco, Microsoft. 2005.

“Improved Lower Bounds for Deletion Channels.” MIT, EPFL (Lausanne). 2005.

“Building a Better Bloom Filter.” Google, Tufts, EPFL (Lausanne), MIT. 2004-05.

“Geometric Generalizations of the Power of Two Choices.” Columbia, Brown, MIT, Georgia Tech. 2004-05.

“New Lower Bounds on the Capacity of the Binary Deletion Channel.” MIT, 2004.

“Digital Fountains, and their Use for Informed Content Delivery across Overlay Networks.” Cornell, 2003.

“Verification Codes for Error and Deletion Channels.” U.C. Berkeley, 2003.

“New Exhaustive, Heuristic, and Interactive Algorithms for 2D Strip Packing.” U.C. Berkeley, 2003.

“Dynamic Models for File Sizes and Double Pareto Distributions.” IBM Research, Microsoft Silicon Valley, Stanford, Tufts, Boston University, Random Graphs and Algorithms Workshop. 2002.

“Verification Codes.” DIMACS workshop. 2001.

“Getting Across Town by Bus, Aggregation, and Reliability.” MIT (LIDS), Mitsubishi Electronic Research Labs. 2001.

“New Ideas for Layered Multicast Congestion Control.” International Computer Science Institute. 2001.

“Compressed Bloom Filters, and Towards Compressing Web Graphs.” U. of Washington, AT&T Labs, U. C. Berkeley, MIT, Brandeis, Carnegie Mellon University, Stanford, DIMACS workshop on compression, U.C. San Diego. 2001.

“Toward Compressing Web Graphs.” Compaq Systems Research Center, Altavista Corporation. 2000.

“Improved Results for Route Planning in Stochastic Transportation Networks.” Compaq Systems Research Center. 2000.

“Choosing the Shortest of Two, Improved, with Applications to IP Routing.” U. Washington, Princeton, Yale, U.C. Berkeley, Boston University, Stanford, M.I.T., Northeastern, Compaq Systems Research Center. 2000.

“Tornado Codes, with Applications to Reliable Multicast.” Harvard, CMU, Duke, AT&T Labs, Univ. of Illinois at Urbana-Champaign, University of Rhode Island, Boston University. 1998-2000.

“How Useful Is Old Information?” MIT, U.C. Berkeley, Stanford, Microsoft. 1997.

“Practical Loss-Resilient Codes.” U.C. Berkeley, Harvard. 1997.

“An Average Case Analysis for First-Fit Bin Packing.” Stanford, DIMACS workshop. 1997.

“The Power of Two Choices.” U.C. Berkeley, Brown, Harvard, Boston University, IBM Almaden, Digital Systems Research Center. 1997.

“The Power of Two Choices and Other Examples of Using Differential Equations to Analyze Algorithms.” U. Mass.