Notice of Approval: New Submission

May 1, 2019

Principal Investigator: Michael Bailey
CC: Deepak Kumar, Kirill Levchenko, Paul Murley, Joshua Mason, Joshua Reynolds

Protocol Title: Building a Map of DNS Resolvers Around the World
Protocol Number: 19739
Funding Source: ICR
Review Type: Exempt 2
Status: Active
Risk Determination: No more than minimal risk
Approval Date: May 1, 2019
Closure Date: April 30, 2024

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is responsible for:

- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Using the approved consent documents, with the footer, from this approved package.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.
Guidelines for completing this research protocol:

- Please submit typed applications via email. Handwritten forms and hard copy forms will not be accepted.
- For items and questions that do not apply to the research, indicate as “not applicable.”
- Provide information for all other items clearly and avoid using discipline-specific jargon.
- Please only include text in the provided boxes. The text boxes will expand as they are typed in to accommodate large amounts of text.
- Ensure that your research qualifies as exempt. Exempt categories of research can be viewed [here](#). If the proposed research does not qualify in any of these categories, please complete and submit the Protocol Form.

Before submitting this application, ensure that the following have been completed.

- Exempt Form is complete.
- Relevant CITI modules have been completed for all members of the research team at [www.citiprogram.org](http://www.citiprogram.org).
- Informed consent/assent/parental permission document(s) are provided.
- Recruitment materials are provided.
- Research materials (e.g. surveys, interview guides, etc.) are provided.
- Any relevant letters of support are provided.

Instructions on the exempt review process and guidance to submitting applications, can be found on the OPRS website. You may also contact OPRS by email at irb@illinois.edu or phone at 217-333-2670.

Submit completed applications via email to: irb@illinois.edu.
Exempt Form

Section 1: PRINCIPAL INVESTIGATOR (PI)

The Illinois Campus Administrative Manual allows assistant, associate, and full professors to act as PI. Other individuals may serve as PI after obtaining approval from the necessary party.

<table>
<thead>
<tr>
<th>Last Name: Bailey</th>
<th>First Name: Michael</th>
<th>Degree(s): PhD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dept. or Unit: ECE</td>
<td>Office Address: Coordinated Science Laboratory 444</td>
<td></td>
</tr>
<tr>
<td>Street Address: 1308 W. Main Street</td>
<td>City: Urbana</td>
<td>State: IL</td>
</tr>
<tr>
<td>Phone: 217-244-8830</td>
<td>E-mail: <a href="mailto:mdbailey@illinois.edu">mdbailey@illinois.edu</a></td>
<td></td>
</tr>
</tbody>
</table>

Urbana-Champaign Campus Status:
- Non-visiting member of (Mark One)
  - Faculty ☑
  - Academic Professional/Staff ☐

(Student Investigators cannot serve as PI)

Training
- ☑ Required CITI Training, Date of Completion (valid within the last 3 years), 2/16/2018
- ☐ Additional training, Date of Completion,

Section 2. RESEARCH TEAM

2A. Are there other investigators engaged in the research?
- ☑ Yes (include a Research Team Form)
- ☐ No

2B. If yes, are any of the researchers not affiliated with Illinois?
- ☐ Yes
- ☑ No

Section 3. PROTOCOL TITLE

Building a Map of DNS Resolvers Around the World

Section 4. FUNDING INFORMATION

4A. Is your research funded? ☑ Yes ☐ No

If no, is there a pending funding decision? ☑ Yes ☐ No

4B. If either of the above were answered yes, please indicate the funding agency: Dr. Michael Bailey ICR (ECE Dept)

4C. A copy of the funding proposal is included: ☑ Yes

Section 5. CONFLICTS OF INTEREST

Please indicate below whether any investigators or members of their immediate families have any of the following. If the answer to any of the following items is yes, please submit the University of Illinois approved conflict management plan. If you have any questions about conflicts of interest, contact coi@illinois.edu.

Office for the Protection of Research Subjects
University of Illinois at Urbana-Champaign
Approved May 1, 2019
IRB #19739
5A. Financial interest or fiduciary relationship with the research sponsor (e.g. investigator is a consultant for the research sponsor). [ ] Yes ☒ No

5B. Financial interest or fiduciary relationship that is related to the research (e.g. investigator owns a startup company, and the intellectual property developed in this protocol may be useful to the company). [ ] Yes ☒ No

5C. Two or more members of the same family are acting as research team members on this protocol. [ ] Yes ☒ No

Section 6. RESEARCH SUMMARY

6A. In lay language, summarize the objective and significance of the research.
Almost all Internet systems rely on the Domain Name System (DNS) to function properly. DNS provides a mapping between human-understandable domain names (e.g., facebook.com.) and IP addresses (e.g., 157.240.2.35), which are series of numbers that uniquely identify the “location” of a remote server. Due to its importance, many researchers have studied and leverages DNS and its properties, from its latency [1], to its growing importance in online censorship [2,3]. To leverage DNS, Internet clients communicate with DNS resolvers – these are computers whose main responsibility is to look up IP addresses to domain names. Unfortunately, there are no up to date studies on the underlying map of the DNS space – that is, what DNS resolver do Internet clients use by default?

In this proposal, we investigate two primary questions about the DNS ecosystem:
1. What DNS resolvers do clients use by default?
2. When clients choose to use a DNS resolver that is out of network, why do they do this? Are there popular custom DNS resolvers that many clients use?

Answering these questions will illuminate the current state of DNS architecture, and will be useful for Internet scientists to build off for future research.

For this project, we will be collecting two pieces of information – the client IP address used to resolve domains and the IP address of the DNS resolver. For the client IP address, we take special precautions to anonymize the last octet of the IP address so it cannot be distinguished from anyone else in the same /24 subnetwork. This offers an anonymity set of 256 users, though in practice the set may be smaller. We do not anonymize the IP address of the DNS resolver – this is a typically shared infrastructure and cannot be tied back to any specific computer.

We note that while DNS data typically contains potentially sensitive information about user behavior (i.e., domain names looked up may reveal political affiliations or medical illnesses[4]), our experiment is specifically designed to only lookup domains in our control and solicit no other information from the Internet client.

To collect the data described, we will use a web advertisement campaign which loads a single domain under our control. The advertisement leverages JavaScript to request the domain under our control. We
note that while JavaScript has the potential to execute dangerous and sometimes malicious code, in this context we are making a single web request to a safe domain that we control. This process is thus identical to any other advertisement a user may see while browsing the web. While our measurements will require diversity in networks (ideally, globally), we will not target individuals or users based on any personal demographic information (e.g., gender, race, political affiliation, etc.)


6B. Indicate if your research includes any of the following:
☐ Secondary data (use of data collected for purposes other than the current research project)
☐ Data collected internationally (include International Research Form)
☐ Translated documents (include Certificate of Translation and translated documents)
☐ Research activities will take place at Carle

6C. Letters of support from outside institutions or entities that are allowing recruitment, research, or record access at their site(s) are attached. ☐ Yes ☑ Not Applicable

Section 7. PARTICIPANTS AND RECRUITMENT

7A. What is the estimated total number of participants? 1 million

7B. Select all participant populations that will be recruited, either intentionally or unintentionally.
Age:
☒ Adults (18+ years old)
☐ Minors (≤17 years old)
☐ Specific age range, please specify:

Gender:
☒ No targeted gender (both men and women will be recruited/included)
☐ Targeted gender, please indicate: ☐ Men/boys ☐ Women/girls ☐ Other, please specify:

Race/Ethnicity:
☒ No targeted race or ethnicity (all races and ethnicities will be recruited/included)
☐ Targeted race or ethnicity, please specify:

College Students:
☒ No targeted college population
☐ UIUC general student body
☐ Targeted UIUC student population, provide the instructor or course information, name of the departmental subject pool, or other specific characteristics:
☐ Students at institution(s) other than UIUC, please specify:

Any research with students on UIUC’s campus needs to be registered with the Office of the Dean of Students.

Other:
People who are illiterate or educational disadvantaged
People who are low-income or economically disadvantaged
People who are non-English speaking
Other, please specify:

7C. Select all recruitment procedures that will be used.
- Student subject pool, please specify:
- Email distribution
- MTurk, Qualtrics Panel, or similar online population
- US Mail
- Flyers
- Website ad, online announcement (e.g. eWeek), or other online recruitment
- Newspaper ad
- Verbal announcement
- Other, please specify: Web HTML ads to collect data
- Not applicable (secondary data only)

Drafts or final copies of all recruitment materials are attached. ☒ Yes

7D. For each group of participants, describe the details of the recruitment process.
We will collect data from users through web advertisements – though participants are not “recruited” in the traditional sense. We will use both the Google Ads platform and the Adroll platform to distribute the web advertisements. The data is collect when the ad loads on the webpage – their computer will make a DNS lookup to a safe domain that we control.

[1] https://ads.google.com/

7E. Will subjects receive compensation or rewards before, during, or after participation?
- Yes ☒ No
If yes, provide a brief description of compensation or rewards.

Section 8. RESEARCH PROCEDURES

8A. Select all research methods and/or data sources that apply.
- Surveys or questionnaires, select all that apply: ☐ Paper ☐ Telephone ☐ Online
- Interviews
- Focus groups
- Field work or ethnography
- Standardized written, oral, or visual tests
- Taste or smell testing
- Intervention or experimental manipulation
- Recording audio and/or video and/or taking photographs
- Materials that have already been collected or already exist, specify source of data:
- HIPAA-protected data
Exempt Form

☐ FERPA-protected data
☒ GDPR-protected data
☐ Other, please specify:

8B. List all testing instruments, surveys, interview guides, etc. that will be used in this research.
- HTML5 advertisement that performs a DNS lookup
Drafts or final copies of all research materials are attached. ☒ Yes

8C. List all locations where research will take place.
The web

8D. List approximate study dates. April 2019 – December 2019

8E. What is the duration of participants’ involvement? One webpage load

8F. How many times will participants engage in research activities? One time

8G. Narratively describe the research procedures in the order in which they will be conducted.
There is no participant involvement in this research procedure. Participants will see a web advertisement that displays our research groups’ logo (attached). This advertisement will make a web request back to a safe domain that we control, and we will log an anonymized version of the client IP address as well as the full, un anonymized IP address of the DNS resolver used to perform the lookup. We will solicit no additional information.

Section 9. CONFIDENTIALITY AND PRIVACY

9A. How are participant data, records, or specimens identified when received or collected by researchers? Identifiers include, but are not limited to, name, date of birth, email address, street address, phone number, audio or video recordings, and SSN.
☐ No identifiers are collected or received
☐ Direct identifiers
☒ Indirect identifiers (e.g. a code or pseudonym used to track participants);
Does the research team have access to the identity key? ☒ Yes ☐ No

9B. Select all methods used to safeguard research records during storage:
☐ Written consent, assent, or parental permission forms are stored separately from the data
☐ Data is collected or given to research team without identifiers
☐ Data is recorded by research team without identifiers
☐ Direct identifiers are removed from collected data as soon as possible
☐ Direct identifiers are deleted and no identity key exists as soon as possible
☒ Participant codes or pseudonyms are used on all data and the existing identity key is stored separately from the data
☐ Electronic data is stored in a secure, UIUC-approved location, please specify
☐ Hard-copy data is stored in a secure location On UIUC’s campus, please specify
☒ Other, please specify: No identity key is collected by the researchers. Data will be stored long-term on password-protected lab servers in the FERPA-compliant, access-controlled UIUC Advanced Computer Building. Data will also be on password-protected computers during analysis

9C. How long will identifiable data be kept? We will not collect no identifiable information.
9D. Describe provisions to protect the privacy interests of subjects. For the client IP address, we take special precautions to anonymize the last octet of the IP address so it cannot be distinguished from anyone else in the same /24 subnetwork. This offers an anonymity set of 256 users, though in practice the set may be smaller. The anonymity set will always 256 IP addresses without the usage of secondary data (e.g., usage patterns of the IP address space). We note that we do not collect or use any secondary data in this study. In addition, we note that depending on how the IP address is being used, there may likely be hundreds or thousands of users behind a single IP address (through Network Address Translation), and so the likelihood of tracing back to an individual in this study is at max 1/256 but may be even smaller.

9E. Describe the training and experience of all persons who will collect or have access to the data. All researchers on this project have extensive experience in Internet science and conducting ethical measurements therein. The responsible primary investigator, Michael Bailey, has published over 60 academic papers, many of which involve human subjects and thus IRB approval. The secondary investigator (Deepak Kumar) is a third-year PhD student who has authored six conference publications. The other investigators are other graduate students, two undergraduate research assistants, two industry user experience researchers, and a research scientist collaborating with Michael Bailey. We intend to make the aggregate data available for other researchers who wish to validate our conclusions.

Section 10. CONSENT PROCESS

10A. Indicate all that apply for the consent process.
- [ ] Written informed consent
- [ ] Waiver of Documentation (signature) of Informed Consent
  - [ ] Online consent  [ ] Oral consent  [ ] Unsigned Information Sheet Provided
  
  Explain why a Waiver of Documentation is necessary:
  
  [ ] Waiver of Informed Consent
  
  Explain why a Waiver of Informed Consent is necessary:
  This research involves no more than minimal risk. The only identifiable information collected from the user (their client IP address) will be anonymized and un-linkable back to a single individual. Based on the mechanism used to deploy the research (i.e., web advertisements), it would additionally be infeasible to request informed consent from any participant without heavily biasing the data to participants who, for example, click on advertisements. In addition, our web advertisements behave no differently than any other web advertisement a user may interact with on the web; it makes only two web requests to safe domains. As the value of this study comes from its scale and effectiveness in measuring the Internet client population, a waiver of informed consent is necessary.

10B. List all researchers who will obtain consent from participants. n/a

10C. Describe the informed consent process. n/a

10D. Where will consent be obtained? n/a

10E. Will participants receive a copy of the consent form for their records?
- [ ] Yes  [ ] No, if no, explain: n/a

University of Illinois at Urbana-Champaign
Institutional Review Board

Approved May 1, 2019
IRB #19739
### 10F. Indicate factors that may interfere or influence the collection of voluntary informed consent.

- No known factors
- Research will involve students enrolled in a course or program taught by a member of the research team
- Research will involve employees whose supervisor(s) is/are recruiting participants
- Participants have a close relationship to the research team
- Other, specify any relationship that exists between the research team and participants:

If applicable, describe the procedures to mitigate the above factors.

### 10G. Copies of the consent form(s) are attached.
- Yes
- Not applicable

---

### Section 11. DISSEMINATION OF RESULTS

11A. List proposed forms of dissemination (e.g. journal articles, thesis, academic paper, conference presentation, sharing within industry, etc.).

Information produced by this project will be disseminated through publication at a top tier computer science conference.

11B. Will any identifiers be published, shared, or otherwise disseminated?
- Yes
- No

If yes, does the consent form explicitly ask consent for such dissemination, or otherwise inform participants that it is required in order to participate in the study?
- Yes

11C. Do you intend to put de-identified data in a data repository?
- Yes
- No

If yes, explain how data will be de-identified. The only potentially identifiable information (client IP address) will be anonymized before dissemination of the data collected.

---

### Section 12. EXPECTED COMPLETION DATE

12A. What is the expected completion date of this research?
- Nov 2019

12B. Please note: Exempt protocols are given a closure date 5 years after their initial approval date, although researchers can request that the study remain open as the closure date approaches.

---

### Section 13 INVESTIGATOR ASSURANCES

- I certify that the project described above, to the best of my knowledge, qualifies as an exempt study.
- I agree that any changes to the project will be submitted to the Office for the Protection of Research Subjects for review prior to implementation. I realize that some changes may alter the exempt status of this project.

The original signature of the PI is required before this application may be processed (electronic signatures are acceptable).
Section 14. DEPARTMENTAL ASSURANCE (OPTIONAL)
If the PI is not eligible to serve as PI under the Campus Administrative Manual, the applicable academic dean, institute director, or campus administrative officer indicates their approval of the researcher to act as Principal Investigator.

<table>
<thead>
<tr>
<th>Principal Investigator</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Bailey</td>
<td>4/29/19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applicable Authorizing Officer</th>
<th>Date</th>
</tr>
</thead>
</table>
For Listing Additional Researchers who are Involved in the Project
All forms must be typewritten and submitted via email to irb@illinois.edu.

When to use this form: If there are collaborating researchers participating in a research study, including those from other institutions, complete this form by listing all collaborating researchers. Include all persons who will be: 1) directly responsible for project oversight and implementation, 2) recruitment, 3) obtaining informed consent, or 4) involved in data collection, analysis of identifiable data, and/or follow-up. Please copy and paste text fields to add additional research team members.
Note:
- Changes made to the Principal Investigator require a revised Protocol Form and an Amendment Form.
- A complete Research Team form with all research team members included needs to be submitted every time the research team is updated.

Section 1. PROTOCOL INFORMATION

1A. Principal Investigator: Michael Bailey
1B. Protocol Number:
1C. Project Title: Building a Map of DNS Resolvers Around the World

Section 2. ADDITIONAL INVESTIGATORS

<table>
<thead>
<tr>
<th>Full Name</th>
<th>Degree</th>
<th>Dept. or Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kirill Levchenko</td>
<td>PhD</td>
<td>ECE</td>
</tr>
<tr>
<td>Professional Email: <a href="mailto:klevchen@illinois.edu">klevchen@illinois.edu</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campus Affiliation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☑️ University of Illinois at Urbana-Champaign</td>
<td>☐ Other, please specify:</td>
<td></td>
</tr>
<tr>
<td>Campus Status:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☑️ Faculty ☐ Academic Professional/Staff ☑️ Graduate Student ☐ Undergraduate Student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☐ Visiting Scholar ☐ Other, please specify:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☑️ Required CITI Training, Date of Completion (valid within last 3 years): 29 April 2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☐ Additional training, Date of Completion:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☑️ This researcher should be copied on OPRS and IRB correspondence.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Full Name</th>
<th>Degree</th>
<th>Dept. or Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joshua Reynolds</td>
<td>B.S.</td>
<td>CS</td>
</tr>
<tr>
<td>Professional Email: <a href="mailto:joshuar3@illinois.edu">joshuar3@illinois.edu</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone: 916-676-6076</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campus Affiliation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☑️ University of Illinois at Urbana-Champaign</td>
<td>☐ Other, please specify:</td>
<td></td>
</tr>
<tr>
<td>Campus Status:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Research Team

<table>
<thead>
<tr>
<th>Faculty</th>
<th>Academic Professional/Staff</th>
<th>Graduate Student</th>
<th>Undergraduate Student</th>
<th>Visiting Scholar</th>
<th>Other, please specify:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Training:**
- Required CITI Training, **Date of Completion** (valid within last 3 years): 14 Sept 2017
- Additional training, **Date of Completion:**
- This researcher should be copied on OPRS and IRB correspondence.

**Full Name:** Joshua Mason  
**Degree:** PhD  
**Dept. or Unit:** ECE  
**Professional Email:** joshm@illinois.edu  
**Phone:**

**Campus Affiliation:**
- University of Illinois at Urbana-Champaign  
- Other, please specify:

**Campus Status:**
- Faculty  
- Academic Professional/Staff  
- Graduate Student  
- Undergraduate Student  
- Visiting Scholar  
- Other, please specify:

**Training:**
- Required CITI Training, **Date of Completion** (valid within last 3 years): Mar 5 2018
- Additional training, **Date of Completion:**
- This researcher should be copied on OPRS and IRB correspondence.

**Full Name:** Deepak Kumar  
**Degree:** B.S.  
**Dept. or Unit:** CS  
**Professional Email:** dkumar11@illinois.edu  
**Phone:** 248-231-3575

**Campus Affiliation:**
- University of Illinois at Urbana-Champaign  
- Other, please specify:

**Campus Status:**
- Faculty  
- Academic Professional/Staff  
- Graduate Student  
- Undergraduate Student  
- Visiting Scholar  
- Other, please specify:

**Training:**
- Required CITI Training, **Date of Completion** (valid within last 3 years): 4/9/19
- Additional training, **Date of Completion:**
- This researcher should be copied on OPRS and IRB correspondence.

**Full Name:** Paul Murley  
**Degree:** M.S.  
**Dept. or Unit:** CS  
**Professional Email:** pmurley2@illinois.edu  
**Phone:** 512-569-4645

**Campus Affiliation:**
- University of Illinois at Urbana-Champaign  
- Other, please specify:

**Campus Status:**
- Faculty  
- Academic Professional/Staff  
- Graduate Student  
- Undergraduate Student  
- Visiting Scholar  
- Other, please specify:

**Training:**
- Required CITI Training, **Date of Completion** (valid within last 3 years): 10/12/17
<table>
<thead>
<tr>
<th></th>
<th>Degree: B.S.</th>
<th>Dept. or Unit: CS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Name:</strong> Zane Ma</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Professional Email:</strong> <a href="mailto:zanema2@illinois.edu">zanema2@illinois.edu</a></td>
<td></td>
<td>Phone: (310)-498-0016</td>
</tr>
<tr>
<td><strong>Campus Affiliation:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Illinois at Urbana-Champaign</td>
<td></td>
<td>Other, please specify:</td>
</tr>
<tr>
<td><strong>Campus Status:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faculty Academic Professional/Staff</td>
<td>Graduate Student</td>
<td>Undergraduate Student</td>
</tr>
<tr>
<td>Visiting Scholar Other, please specify:</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Training:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Required CITI Training, Date of Completion (valid within last 3 years): 2/9/19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional training, Date of Completion:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>This researcher should be copied on OPRS and IRB correspondence.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>