COMPUTER SCIENCE HONORS & RESEARCH

DAVE LEVIN

ASST PROF & CHAIR OF CS HONORS

COMPUTER SCIENCE HONORS

Departmental honors program

GOALS

Give you opportunities to
try out research

Set you up to get into a great
grad school (if you so choose)

Help create a network of you so
you can ***do great things*** together

COMPUTER SCIENCE RESEARCH

WHAT IS CS RESEARCH?

CS	The study and application of computing
+ **RESEARCH**	Asking and answering unsolved questions
CS RESEARCH	Doing things with computers that no one has done before

SOME EARLY CS RESEARCH QUESTIONS

What is ***possible*** to compute?
What is ***efficient*** to compute?

→ *P = NP?*

Big data ←

How much data can we ***store***?
How much data can we ***search***?

How can computers ***communicate***?
How can computers ***parallelize***?
How can we ***program*** networks?

→ *Cloud computing*

WHAT KIND OF RESEARCH IS THIS?

What is the structure of ***social networks***?

Which articles are ***"fake news"***?

Is it possible to ***see through walls***?

Can we predict a movie's opening ***box office revenue***?

How do people ***watch TV***?

Can we protect ***free speech*** in oppressive regimes?

THESE ARE ALL CS PROBLEMS, TOO!

WHAT IS THE STRUCTURE OF SOCIAL NETWORKS?

Measurement and Analysis of Online Social Networks

Alan Mislove
MPI for Software Systems
Campus E1 4
Saarbrücken 66123, Germany

Massimiliano Marcon
MPI for Software Systems
Campus E1 4
Saarbrücken 66123, Germany

Krishna P. Gummadi
MPI for Software Systems
Campus E1 4
Saarbrücken 66123, Germany

Peter Druschel
MPI for Software Systems
Campus E1 4
Saarbrücken 66123, Germany

Bobby Bhattacharjee
Computer Science Department
University of Maryland
College Park, MD 20742

ABSTRACT

Online social networking sites like Orkut, YouTube, and Flickr are among the most popular sites on the Internet. Users of these sites form a social network, which provides a powerful means of sharing, organizing, and finding content and contacts. The popularity of these sites provides an opportunity to study the characteristics of online social network graphs at large scale. Understanding these graphs is important, both to improve current systems and to design new applications of online social networks.

This paper presents a large-scale measurement study and analysis of the structure of multiple online social networks. We examine data gathered from four popular online social networks: Flickr, YouTube, LiveJournal, and Orkut. We crawled the publicly accessible user links on each site, obtaining a large portion of each social network's graph. Our data set contains over 11.3 million users and 328 million links. We believe that this is the first study to examine multiple online social networks at scale.

Our results confirm the power-law, small-world, and scale-free properties of online social networks. We observe that the indegree of user nodes tends to match the outdegree; that the networks contain a densely connected core of high-degree nodes; and that this core links small groups of strongly clustered, low-degree nodes at the fringes of the network. Finally, we discuss the implications of these structural properties for the design of social network based systems.

Categories and Subject Descriptors

H.5.m [Information Interfaces and Presentation]: Miscellaneous; H.3.5 [Information Storage and Retrieval]: Online Information Services—*Web-based services*

General Terms

Measurement

Keywords

Social networks, measurement, analysis

1. INTRODUCTION

The Internet has spawned different types of information sharing systems, including the Web. Recently, *online social networks* have gained significant popularity and are now among the most popular sites on the Web [40]. For example, MySpace (over 190 million users[1]), Orkut (over 62 million), LinkedIn (over 11 million), and LiveJournal (over 5.5 million) are popular sites built on social networks.

Unlike the Web, which is largely organized around content, online social networks are organized around users. Participating users join a network, publish their profile and any content, and create links to any other users with whom they associate. The resulting social network provides a basis for maintaining social relationships, for finding users with similar interests, and for locating content and knowledge that has been contributed or endorsed by other users.

An in-depth understanding of the graph structure of online social networks is necessary to evaluate current systems, to design future online social network based systems, and to understand the impact of online social networks on the Internet. For example, understanding the structure of online social networks might lead to algorithms that can detect trusted or influential users, much like the study of the Web graph led to the discovery of algorithms for finding authoritative sources in the Web [21]. Moreover, recent work has proposed the use of social networks to mitigate email spam [17], to improve Internet search [35], and to defend against Sybil attacks [55]. However, these systems have not yet been evaluated on real social networks at scale, and little is known to date on how to synthesize realistic social network graphs.

In this paper, we present a large-scale (11.3 million users, 328 million links) measurement study and analysis of the structure of four popular online social networks: Flickr, YouTube, LiveJournal, and Orkut. Data gathered from multiple sites enables us to identify common structural properties of online social networks. We believe that ours is the first study to examine multiple online social networks at scale. We obtained our data by crawling publicly accessible information on these sites, and we make the data available

[1]Number of distinct identities as reported by the respective sites in July 2007.

Most social networks are scale-free graphs (hub-like core)

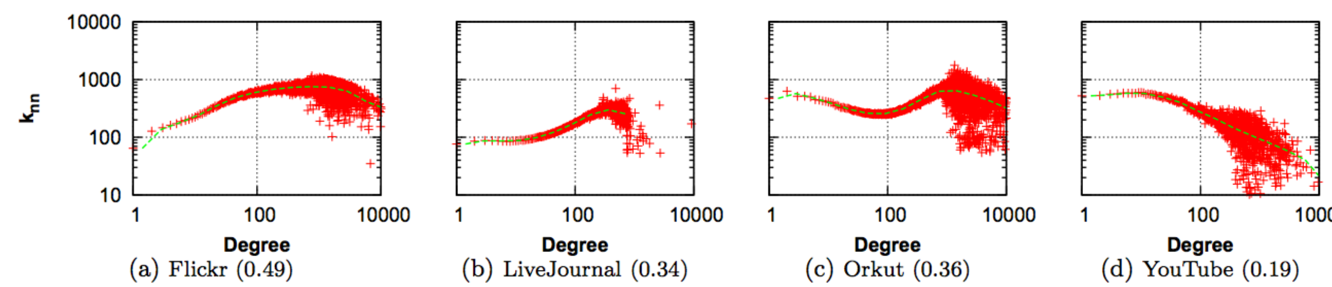


Figure 6: Log-log plot of the outdegree versus the average indegree of friends. The scale-free metrics, included in the legend, suggest the presence of a well-connected core.

CS can answer this in a way sociology could not: ***at scale***

	Flickr
Number of users	1,846,198
Estimated fraction of user population crawled	26.9%
Dates of crawl	Jan 9, 2007
Number of friend links	22,613,981
Average number of friends per user	12.24
Fraction of links symmetric	62.0%
Number of user groups	103,648
Average number of groups memberships per user	4.62

IS IT POSSIBLE TO SEE THROUGH WALLS?

See Through Walls with Wi-Fi!

Fadel Adib and Dina Katabi
Massachusetts Institute of Technology
{fadel,dk}@mit.edu

ABSTRACT

Wi-Fi signals are typically information carriers between a transmitter and a receiver. In this paper, we show that Wi-Fi can also extend our senses, enabling us to see moving objects through walls and behind closed doors. In particular, we can use such signals to identify the number of people in a closed room and their relative locations. We can also identify simple gestures made behind a wall, and combine a sequence of gestures to communicate messages to a wireless receiver without carrying any transmitting device. The paper introduces two main innovations. First, it shows how one can use MIMO interference nulling to eliminate reflections off static objects and focus the receiver on a moving target. Second, it shows how one can track a human by treating the motion of a human body as an antenna array and tracking the resulting RF beam. We demonstrate the validity of our design by building it into USRP software radios and testing it in office buildings.

Categories and Subject Descriptors C.2.2 **[Computer Systems Organization]**: Computer-Communications Networks. H.5.2 **[Information Interfaces and Presentation]**: User Interfaces - Input devices and strategies.

Keywords Seeing Through Walls, Wireless, MIMO, Gesture-Based User Interface

1. INTRODUCTION

Can Wi-Fi signals enable us to see through walls? For many years humans have fantasized about X-ray vision and played with the concept in comic books and sci-fi movies. This paper explores the potential of using Wi-Fi signals and recent advances in MIMO communications to build a device that can capture the motion of humans behind a wall and in closed rooms. Law enforcement personnel can use the device to avoid walking into an ambush, and minimize casualties in standoffs and hostage situations. Emergency responders can use it to see through rubble and collapsed structures. Ordinary users can leverage the device for gaming, intrusion detection, privacy-enhanced monitoring of children and elderly, or personal security when stepping into dark alleys and unknown places.

The concept underlying seeing through opaque obstacles is similar to radar and sonar imaging. Specifically, when faced with a non-metallic wall, a fraction of the RF signal would traverse the wall, reflect off objects and humans, and come back imprinted with a signature of what is inside a closed room. By capturing these reflections, we can image objects behind a wall. Building a device that can capture such reflections, however, is difficult because the signal power after traversing the wall twice (in and out of the room) is reduced by three to five orders of magnitude [11]. Even more challenging are the reflections from the wall itself, which are much stronger than the reflections from objects inside the room [11, 27]. Reflections off the wall overwhelm the receiver's analog to digital converter (ADC), preventing it from registering the minute variations due to reflections from objects behind the wall. This behavior is called the "Flash Effect" since it is analogous to how a mirror in front of a camera reflects the camera's flash and prevents it from capturing objects in the scene.

So how can one overcome these difficulties? The radar community has been investigating these issues, and has recently introduced a few ultra-wideband systems that can detect humans moving behind a wall, and show them as blobs moving in a dim background [27, 41] (see the video at [6] for a reference). Today's state-of-the-art system requires 2 GHz of bandwidth, a large power source, and an 8-foot long antenna array (2.4 meters) [12, 27]. Apart from the bulkiness of the device, blasting power in such a wide spectrum is infeasible for entities other than the military. The requirement for multi-GHz transmission is at the heart of how these systems work: they separate reflections off the wall from reflections from the objects behind the wall based on their arrival time, and hence need to identify sub-nanosecond delays (i.e., multi-GHz bandwidth) to filter the flash effect.[1] To address these limitations, an initial attempt was made in 2012 to use Wi-Fi to see through a wall [13]. However, to mitigate the flash effect, this past proposal needs to install an additional receiver behind the wall, and connect the receivers behind and in front of the wall to a joint clock via wires [13].

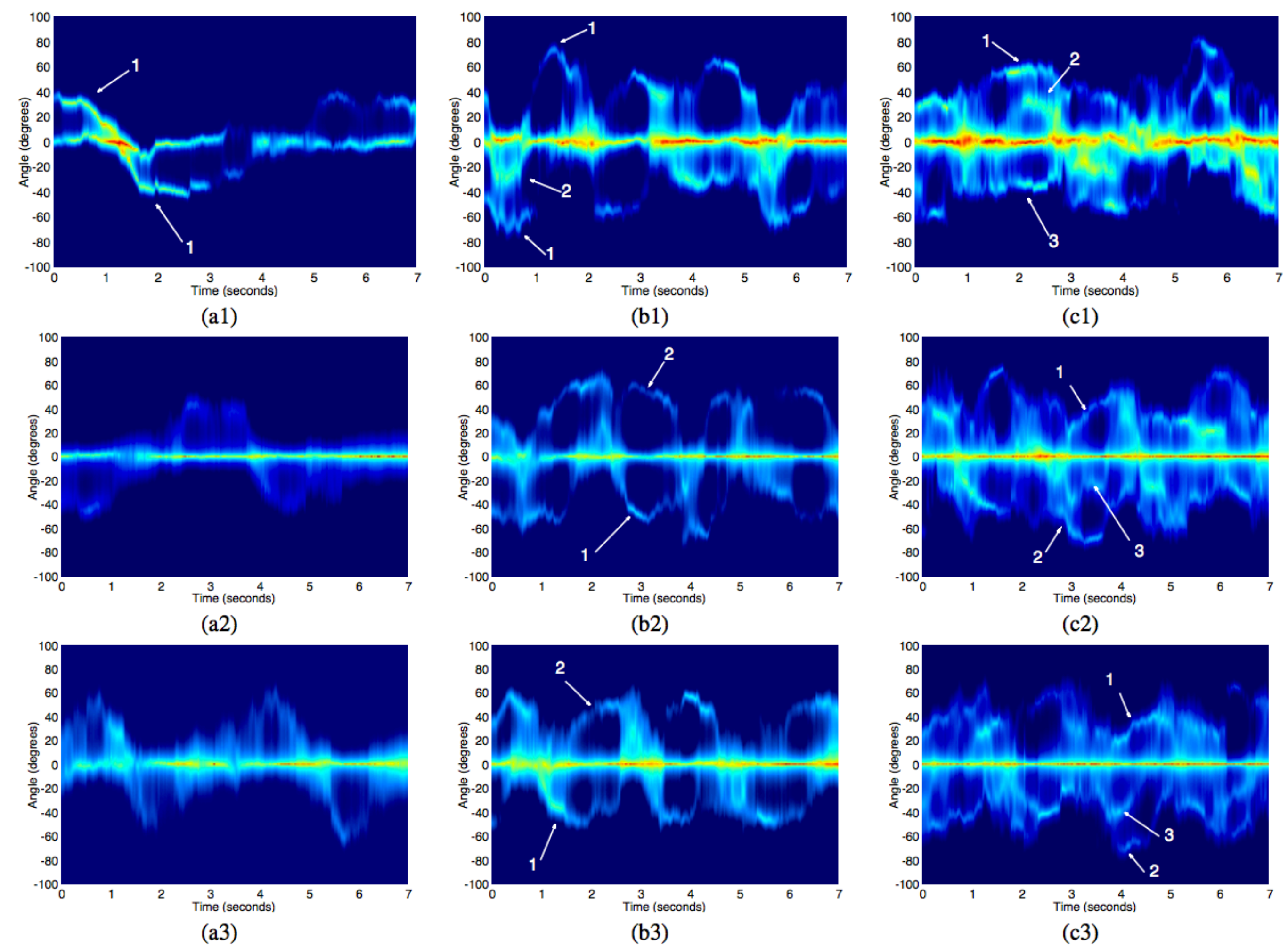
The objective of this paper is to enable a see-through-wall technology that is low-bandwidth, low-power, compact, and accessible to non-military entities. To this end, the paper introduces Wi-Vi,[2] a see-through-wall device that employs Wi-Fi signals in the 2.4 GHz ISM band. Wi-Vi limits itself to a 20 MHz-wide Wi-Fi channel, and avoids ultra-wideband solutions used today to address the flash effect. It also disposes of the large antenna array, typical in past systems, and uses instead a smaller 3-antenna MIMO radio.

So, how does Wi-Vi eliminate the flash effect without using GHz of bandwidth? We observe that we can adapt recent advances in MIMO communications to through-wall imaging. In MIMO, multiple antenna systems can encode their transmissions so that the signal is nulled (i.e., sums up to zero) at a particular receive antenna. MIMO systems use this capability to eliminate interference to unwanted receivers. In contrast, we use nulling to eliminate reflections from static objects, including the wall. Specifically, a Wi-Vi device has two transmit antennas and a single receive antenna. Wi-Vi operates in two stages. In the first stage, it measures the channels from each of its two transmit antennas to its receive antenna. In stage 2, the two transmit antennas use the channel measurements from stage 1 to null the signal at the receive antenna. Since wireless signals (including reflections) combine linearly over the medium, only reflec-

SIGCOMM'13, August 12–16, 2013, Hong Kong, China.

[1] Filtering is done in the analog domain before the signal reaches the ADC.
[2] Wi-Vi stands for Wi-Fi Vision.

Yes, by treating it as a wireless networking problem



(a) One Human (b) Two Humans (c) Three Humans

Figure 8—Tracking human motion with Wi-Vi. The figures show output traces with a different number of humans after processing with the smoothed MUSIC algorithm. They plot $A'[\theta, n]$ where θ is the angle in $[-90, 90]$ is plotted on the y-axis and time is on the x-axis. (a) shows traces for one human; (b) for two humans; and (c) for three humans moving behind the wall of a closed room.

CAN WE PROTECT FREE SPEECH?

Alibi Routing

Dave Levin* Youndo Lee* Luke Valenta† Zhihao Li* Victoria Lai*
Cristian Lumezanu‡ Neil Spring* Bobby Bhattacharjee*

* University of Maryland † University of Pennsylvania ‡ NEC Labs

ABSTRACT

There are several mechanisms by which users can gain insight into where their packets have gone, but no mechanisms allow users undeniable proof that their packets did *not* traverse certain parts of the world while on their way to or from another host. This paper introduces the problem of finding "proofs of avoidance": evidence that the paths taken by a packet and its response avoided a user-specified set of "forbidden" geographic regions. Proving that something did *not* happen is often intractable, but we demonstrate a low-overhead proof structure built around the idea of what we call "alibis": relays with particular timing constraints that, when upheld, would make it impossible to traverse both the relay and the forbidden regions.

We present *Alibi Routing*, a peer-to-peer overlay routing system for finding alibis securely and efficiently. One of the primary distinguishing characteristics of Alibi Routing is that it does not require knowledge of—or modifications to—the Internet's routing hardware or policies. Rather, Alibi Routing is able to derive its proofs of avoidance from user-provided GPS coordinates and speed of light propagation delays. Using a PlanetLab deployment and larger-scale simulations, we evaluate Alibi Routing to demonstrate that many source-destination pairs can avoid countries of their choosing with little latency inflation. We also identify when Alibi Routing does not work: it has difficulty avoiding regions that users are very close to (or, of course, inside of).

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols; C.2.0 [**Computer-Communication Networks**]: General—*Security and protection*

SIGCOMM '15, August 17–21, 2015, London, United Kingdom

 ISBN 978-1-4503-3542-3/15/08…$15.00

DOI: http://dx.doi.org/10.1145/2785956.2787509

Keywords

Alibi Routing; Provable route avoidance; Censorship avoidance; Peer-to-peer; Overlay routing

1. INTRODUCTION

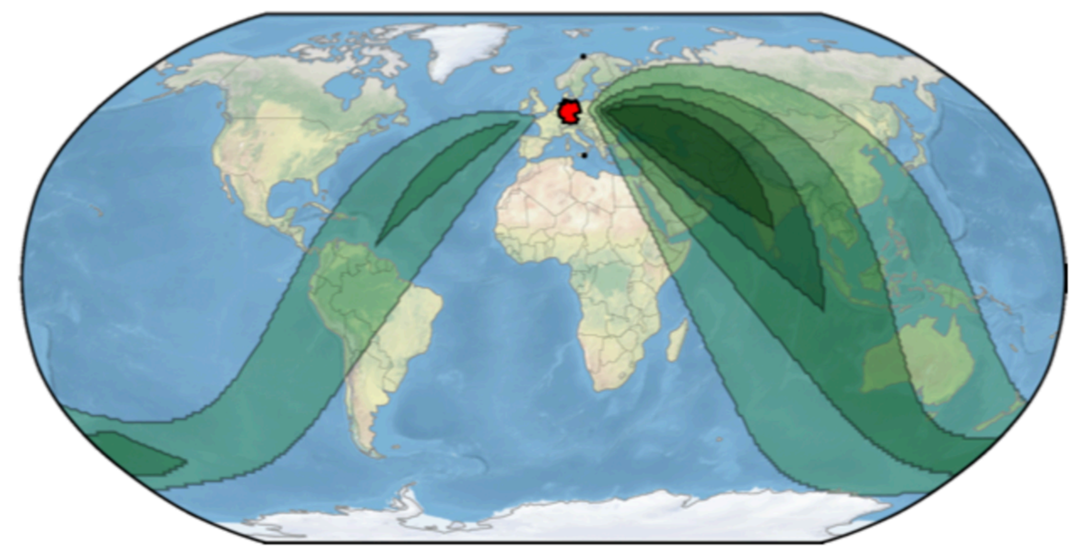
Users have little control over where in the world their packets travel en route to their destinations. Some mechanisms exist to provide insight into where packets traveled, such as the record-route IP option, overlay routing systems (§7), or to a lesser extent source-routing. While these approaches expose a subset of the path the user's packets took, they do not allow a user to determine or provably influence where their packets do *not* go.

This paper introduces a new primitive we call *provable avoidance routing*. With provable avoidance routing, a user specifies arbitrary geographic regions—such as countries or UN voting blocs—to be avoided while communicating with a destination. If successful, the primitive returns *proof* that the user's packets did not traverse the forbidden regions. If it is unsuccessful, it concludes only that the packets *may have* traversed them.

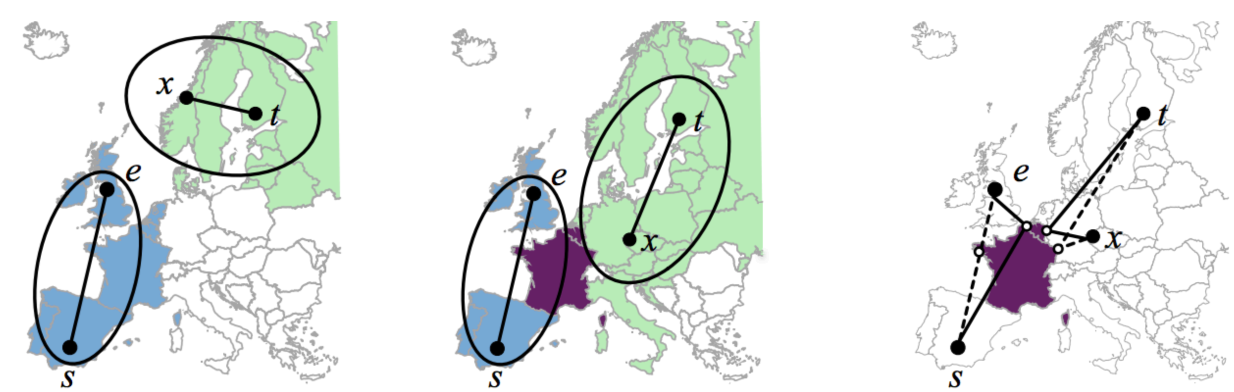
The goal of provable avoidance routing is *detection*, as opposed to *prevention*. In other words, alone, it is unable to ensure a user's packets *will not* traverse a region of the world—we do not require modifications to the underlying routing protocols or hardware, and so we are subject to all of today's uncertainties as to where packets will travel. Rather, what we are able to provide is assurance that the user's packets and their respective responses took paths that *did not* traverse regions of the world. Our proofs of avoidance are provided on a per-packet basis, and are *a posteriori*: only after sending the packet and getting a reply can we ascertain whether or not the round-trip communication avoided the forbidden region.

While outright prevention would be ideal, detection can be a powerful tool, as well. For example, consider one of the greatest threats to open communication on the Internet: censorship. Beyond just dropping [34] or logging [29] users' traffic, censorship can take many forms, including *injecting* packets with false information [4]. Recent results indicate that many users may be censored not by their (or their destination's) countries, but by regimes through which their packets transit; a group of anonymous researchers demonstrated that DNS queries that merely traverse China's borders are

With peer-to-peer networking, route around online censors



Avoid countries that attack Tor



(a) When C_e (in blue) and C_x (in lighter green) do not intersect, double-traversal of any country is impossible.

(b) When C_e and C_x *do* intersect (dark purple), we must compute the shortest distances for both legs to go through each country (right). The dashed lines show the shortest distances through France, and the solid lines through Belgium.

SOME RESEARCH FROM UMD

HAL DAUMÉ III	Can computers ***learn language*** through interaction?
MARINE CARPUAT	Can computers be ***multilingual***?
BOBBY BHATTACHARJEE	Can software extend a phone's ***battery life***?
FOSTER, HICKS, DAVID VAN HORN	How do we ***build software*** that is secure and correct?
YIANNIS ALOIMONOS	Can robots ***learn to cook*** by watching?

SOME RESEARCH FROM UMD

TOM GOLDSTEIN
How can we ***optimize*** machine learning and cloud computing?

ZIA KHAN
How does ***genetic variation*** lead to human-specific traits?

NEIL SPRING
How does ***weather*** affect Internet connectivity?

MICHELLE MAZUREK
What are the ***human factors*** of security?

JOHN DICKERSON
Can better algorithms lead to more efficient ***kidney exchanges***?

TRADITIONAL AREAS OF CS RESEARCH

Artificial Intelligence
AR & VR
Big Data
Comp-bio
Computer systems
Computer vision
Data science
Databases
Graphics
High-perf. computing
HCI
Numerical analysis
Programming languages
Scientific computing
Security
Software engineering
Theory

WHERE IS CS RESEARCH GOING?

Many areas are converging on users

PL
Can we make a language with feature X…? ← Originally
What are the right features to ***help users***…? ← Recently

GFX
Can we realistically render Y…?
How can we make a user experience more ***immersive***?

NET
Can we make a protocol resilient to failure?
Can we make a protocol resilient to ***manipulation***?

DB
Can we store and retrieve data efficiently?
Can we help users ***understand*** data provenance?

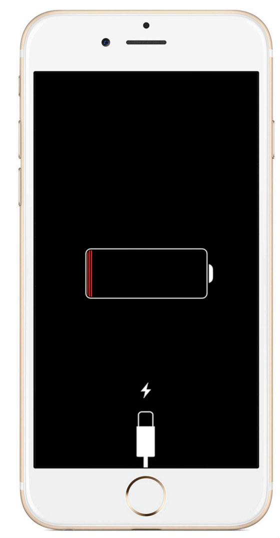
WHERE IS CS RESEARCH GOING?

Computing is becoming ubiquitous

IOT

MOBILE

UBER Instagram



CARS

INSTEAD OF ASKING

WHERE IS CS RESEARCH GOING?

ASK YOURSELF

WHAT ARE THE PROBLEMS THAT NEED SOLVING?

THEN TAKE CS RESEARCH THERE.

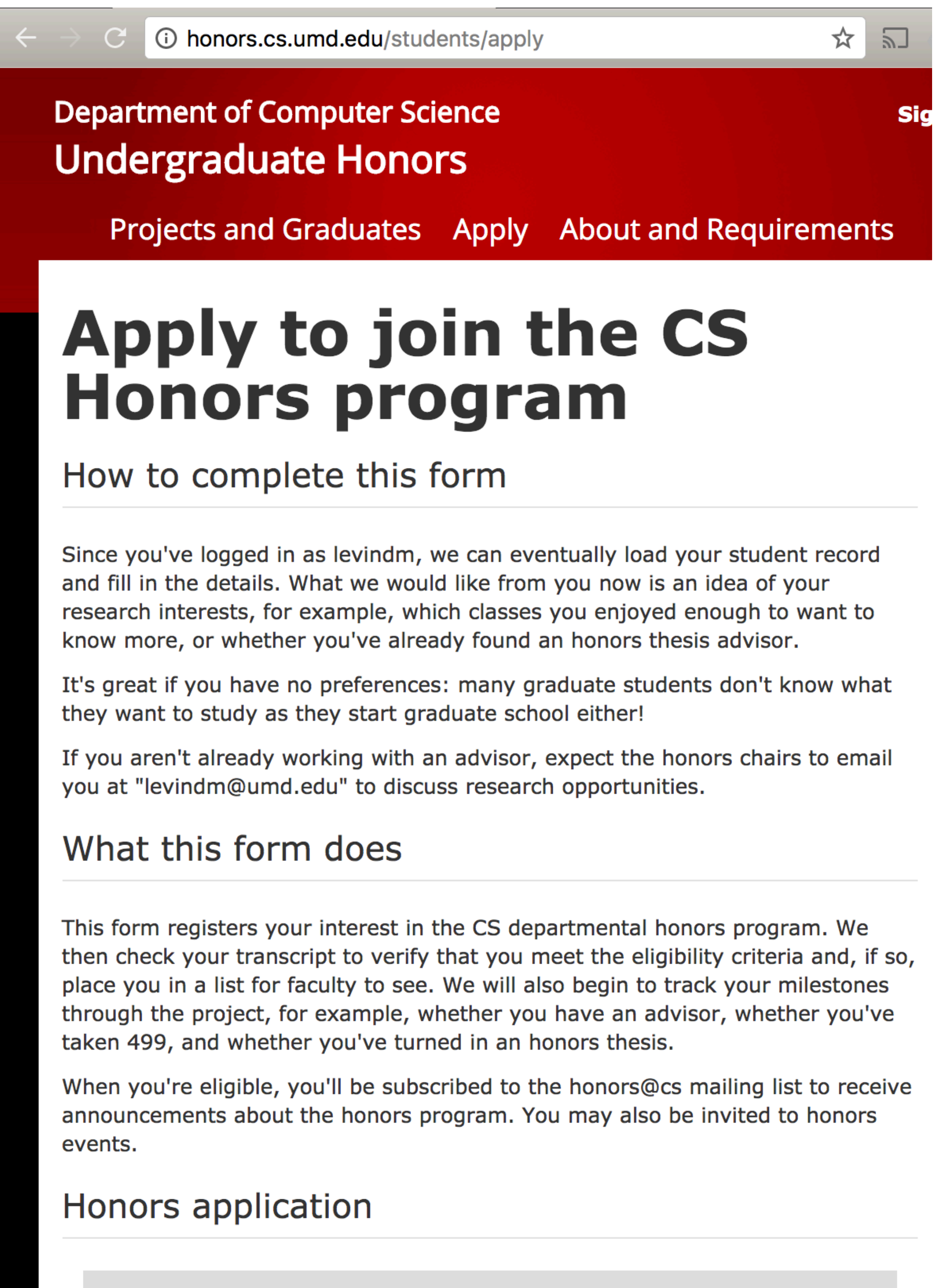
THE FUTURE OF CS IS INTERDISCIPLINARY.

… Economics Law Music Language Math Sociology …

CS HONORS
APPLYING AND GRADUATING

APPLYING TO CS HONORS

HONORS.CS.UMD.EDU



Describe your ***research interests***

- Many grad students aren't 100% sure what their research interests are!
- Just talk about what excites you

You do NOT need to propose a specific research topic

- Identifying a good research problem is the last thing you learn in a PhD

GRADUATING WITH CS HONORS

1	Maintain a good GPA	≥3.50 CS & ≥3.25 All
2	Take an advanced class	H-option or Grad class
3	Take the Honors Seminar	CMSC 396H
4	Do research with faculty	CMSC 499
5	Write an honors thesis	Your contribution

CMSC 396H : HONORS SEMINAR

One-credit class to introduce you to CS research and to faculty researchers

What is research?

How do you perform research effectively and ethically?

What kind of research do our CMSC faculty do?

Gain some experience doing a small research project.

CMSC 499 : INDEPENDENT STUDY

Typically 3-credit class in which you do research with a faculty member

Test the waters of research!

You need to initiate this by talking to the professors

CMSC 396H is a great first step; so is taking the prof's 400/600/700s

What ultimately leads to publications and recommendation letters

LET'S GO TO
GRADUATE SCHOOL

WHY GRAD SCHOOL?

What does having a PhD mean?

THE ABILITY TO

- **Identify** a new problem
- **Solve** a problem in a novel way
- **Evaluate** the solution
- **Communicate** convincingly:
  - why the problem matters
  - why your solution is better

= Intellectual freedom that's needed at **Startups** and in **Academia**

HOW TO GET INTO A GREAT GRAD SCHOOL?

BARE MINIMUM	Good grades; Good GRE scores
GOOD CHANCES	Strong **recommendation letters** from research-active faculty, about research
GREAT CHANCES	**Publication** in a reputable venue (Extra bonus as first author)
ROCK STAR	Give an excellent presentation of the work at the conference

WHEN SHOULD YOU PREP FOR GRAD SCHOOL?

Junior

Fall/Spring

Find a prof to work with

Come up to speed

Perform your research

Summer

Finish paper

Senior

Sep

Submit paper

Optimistic

Oct

Apply for NSF GRFP

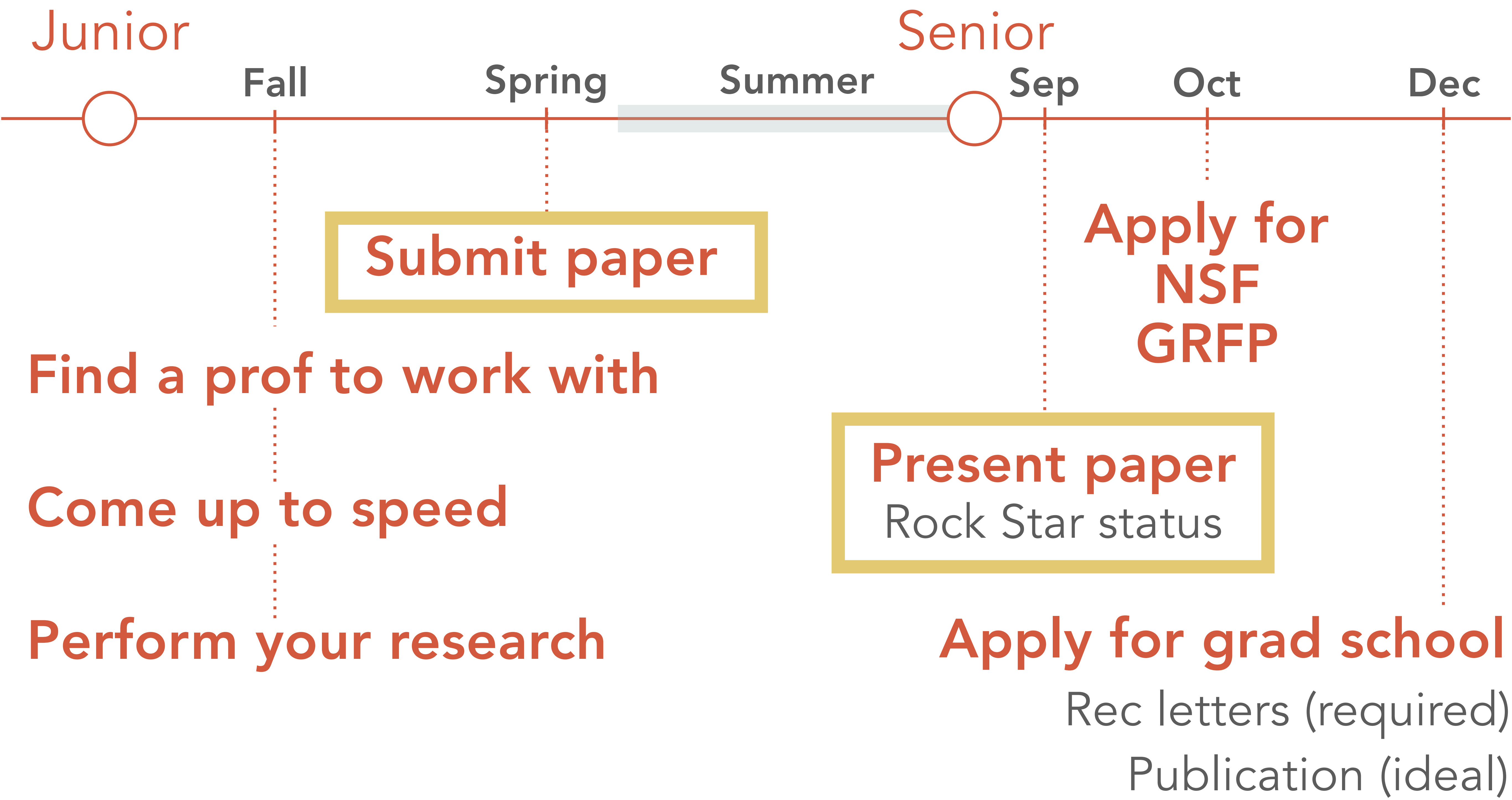
Dec

Apply for grad school

Rec letters (required)

Publication (ideal)

WHEN SHOULD YOU PREP FOR GRAD SCHOOL?



CS HONORS IS ABOUT
GETTING YOU TO EXPLORE NEW OPTIONS

TRY RESEARCH (YOU MIGHT LOVE IT)

DO IT EARLY ENOUGH
TO NOT CLOSE THE DOOR ON A CAREER IN RESEARCH

SEEK OUT OTHER CHALLENGES, TOO
. . ., STARTUP SHELL, SANDBOX, CSEC, . . .

MEET YOUR FELLOW STUDENTS

KATURA HARVEY
SYSTEMS SECURITY

EMILY KOWALCZYK
SOFTWARE TESTING

MATT MYERS
COMPUTATIONAL BIOLOGY

RAMA PADMANABHAN
INTERNET MEASUREMENT

IAN SWEET
PROGRAMMING LANGUAGES

CHALLENGE YOURSELF WITH RESEARCH & CS HONORS

HONORS.CS.UMD.EDU

DAVE LEVIN

ASST PROF & CHAIR OF CS HONORS

DML@CS.UMD.EDU