Cross-Cohort Research Experience for Project Management and Leadership Development

Dr. Yung-Hsiang Lu, Purdue University

Yung-Hsiang Lu is an associate professor in the School of Electrical and Computer Engineering and (by courtesy) the Department of Computer Science of Purdue University. He is an ACM distinguished scientist and ACM distinguished speaker. He is a member in the organizing committee of the IEEE Rebooting Computing Initiative. He is the lead organizer of the first Low-Power Image Recognition Challenge in 2015, the chair (2014-2016) of the Multimedia Communication Systems Interest Group in IEEE Multimedia Communications Technical Committee. He obtained the Ph.D. from the Department of Electrical Engineering at Stanford University.

Prof. Thomas J. Hacker, Purdue University, West Lafayette

Thomas J. Hacker is an Associate Professor of Computer and Information Technology at Purdue University in West Lafayette, Indiana. His research interests include cyberinfrastructure systems, high performance computing, and the reliability of large-scale supercomputing systems. He holds a PhD in Computer Science and Engineering from the University of Michigan, Ann Arbor. He is a member of IEEE, the ACM, and ASEE.

Dr. Carla B. Zoltowski, Purdue University, West Lafayette

Carla B. Zoltowski, Ph.D., is Co-Director of the EPICS Program at Purdue University. She received her B.S. and M.S. in electrical engineering and Ph.D. in engineering education, all from Purdue University. She has served as a lecturer in Purdue's School of Electrical and Computer Engineering. Dr. Zoltowski's academic and research interests broadly include the professional formation of engineers and diversity and inclusion in engineering, with specific interests in human-centered design, engineering ethics, leadership, service-learning, assistive-technology, and accessibility.

Prof. Jan P. Allebach, Purdue University, West Lafayette

Jan P. Allebach is Hewlett-Packard Distinguished Professor of Electrical and Computer Engineering at Purdue University. Allebach is a Fellow of the IEEE, the National Academy of Inventors, the Society for Imaging Science and Technology (IS&T), and SPIE. He was named Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named Honorary Member of IS&T, the highest award that IS&T bestows. He has received the IEEE Daniel E. Noble Award, and is a member of the National Academy of Engineering. He currently serves as an IEEE Signal Processing Society Distinguished Lecturer (2016-2017).

Cross-Cohort Research Experience for Project Management and Leadership Development

Abstract

Project management and leadership skills are essential for career development. However, in typical university settings, undergraduate students take different courses and work on different projects in different teams each semester. As a result, students lack opportunities to work on multi-year projects and develop the skills essential for long-term planning. To remedy this situation, our department has created elective courses that allow students from all years (first-year students to graduate students) to work on research projects under the supervision of faculty members and the mentorship of senior graduate students. These projects provide the opportunities for students to learn many skills essential in workplace, such as (1) understanding how projects are designed and managed; (2) taking responsibilities on different components in the projects; (3) learning computer tools for collaboration and integration; (4) developing leadership skills; (5) cultivating self-learning; and (6) improving communication, both speaking and writing. This paper reports one project that involves undergraduate, masters, and doctoral students. The project, now in its fifth-year, builds computer tools for researchers, educators, and students using cloud computing for large-scale image analysis. The project has received an award in a student competition, and three research grants for international collaboration, entrepreneurship, and big data analytics, and produced more than a dozen research papers. This paper describes the project in detail and shares experiences on many crucial factors necessary for creating a successful cross-cohort research project.

Research experience is not required in typical undergraduate curricula; thus, it is essential to recruit well qualified and interested students. From the beginning of this project, there was a clear goal to create software tools that would become available to the research community. The opportunity to serve real users is appealing to many students. In order to build software tools for users, this project has established rigorous procedures common in commercial software development such as version control, testing, documentation, and so on. Leadership development is another key component: if a student continues in this project over multiple semesters, the student may be promoted to lead a subteam or the entire team. In addition to learning technical skills, the team has participated in multiple student competitions and has won the second prize in one competition. This project also encourages entrepreneurship: a group of students plan to start a company after they have interviewed potential customers exploring the feasibility of commercializing the technology and investigating market-product match. Four foreign institutions are collaborators of the project and the students have experience working with these collaborators through video conferencing.

Introduction

Undergraduate curricula are usually designed for students to explore different areas, including technologies, social science, communication skills, leadership, entrepreneurship, etc. A typical curriculum requires that a student take several courses each semester (or quarter). Other than prerequisites, courses in different semesters are usually independent. Even for courses in a sequence, the connections in different semesters are often informal and may vary based on the experience and interest of the instructors. While “decoupling” among courses serves the purpose for students to explore new areas of knowledge and encourages students to have fresh start every semester, these curricula also have some significant drawbacks, for example,

- A course project can last at most 10 weeks (for a quarter system) to 15 weeks (for a semester system). Students lack the experience gained from participating in and conducting projects that last multiple years. As a result, students tend to develop the habit of short-term thinking and tend to accumulate “technical debts”[1-3]. Technical debts refer to decisions that may be acceptable for the short term, but will eventually need to be repaid. If a course project lasts for only weeks, the debt may be forgiven at the end of the quarter (or semester) and never repaid. When students enter the workplace, they may find it difficult to adjust to a new environment that requires in-depth planning for multi-year projects.
- Students in the same cohort usually take the same courses, without knowing the students in other cohorts. Currently, this is partially remedied by student clubs where different cohorts can meet. Students join clubs voluntarily and may leave clubs at any time. Also, publications and presentations are usually not priorities in student clubs, and thus clubs do not provide the same experience as research projects supervised by faculty members.
- Classroom settings are often different from the “real world”. In “real world”, it is uncommon to have a group of people around the same age, with similar levels of knowledge and experience, working on the same problems (homework assignments) with short deadlines. “Real world” projects usually include people of different levels of experience and each member in the team works on different parts of project. Moreover, “real world” projects have no instructors or teaching assistants who are supposed to know the right answers of homework assignments. Students become accustomed, sometimes with encouragement from the teaching staff, to seek help during office hours. However, there is no office hour in the “real world”.
- Faculty members may have “new and crazy” ideas that may be too risky for graduate students who aim to finish their degrees within two (for MS) or five (for PhD) years. The faculty members may want to form a team and perform preliminary studies without committing significant amounts of resources (such as research assistantships).

It is important understanding the distinction of “student thinking” and “real-world thinking”. Students operate in the acquisition mode: their main goal is to acquire knowledge and skills. They are evaluated by homework assignments and exam questions, both are constrained by the durations of quarters and semesters. Students usually aim to have high GPAs and thus must balance among many courses, i.e., breadth. In contrast, “real world” projects aim to solve problems that are usually ill-defined. There is no weekly homework assignments set by instructors; instead, people must set milestones toward the final goal. To find a good solution, one must have deep knowledge about the problem and existing solutions; hence, the priority is depth, not breadth. Table 1 compares some differences between classroom settings and “real-world” environment.

Table 1 Differences between Typical Classroom Settings and Real World

	Classroom Settings	**Real World**
Project duration	Weeks	Years
Main goal	Acquire knowledge and skills	Solve problems

Problems' solutions	Professors know the solutions	Nobody knows the solution
Team	Same cohort	Diverse
Evaluation methods	Homework assignments and exams	Demonstration, presentations, products
Design decisions	Short term (for one semester)	Long term (for projects' durations)

Purdue University has a history of offering new learning experiences that allow students to participate in multi-year projects. Some of these models have become national and international models[4-6]. One example is a service-learning program, EPICS (Engineering Projects in Community Service), designed for undergraduate students to work with nonprofit organizations and apply knowledge and skills to help communities. Another program is research-oriented, the Vertically Integrated Projects program, and it allows students from all years (first-year to graduate) to participate in research projects supervised by faculty members. VIP teams provide opportunities for students to gain first-hand experience in research projects. VIP teams are different from SURF (summer undergraduate research fellowship) programs[7] available in many universities because SURF projects are limited to summers (even though some universities now provide long-term SURF) and only undergraduate students.

Both the EPICS and VIP programs are multidisciplinary and vertically-integrated (i.e., cross-cohort, first-year through seniors and beyond on the same team). The innovative curricular structure that support both programs allows students to participate for multiple semesters. The programs have also addressed the curricular challenges of multidisciplinary and diverse teams, with robust curricular and assessment structures established to support the service-learning and research activities. Both formal research and informal feedback indicate that these programs are very effective in developing professional skills such as teamwork, leadership, and communication, presenting the discipline in context, and filling a gap between their technical courses and expectations of practice and research[8]. VIP has been adopted by multiple universities as a model for STEM education[9].

This paper describes one VIP project that is focused on the development of a cloud-based software system for large-scale analysis of visual data (image video). More than 50 undergraduate students have contributed to this project. This system is designed to be a tool available to researchers and more than 200 users have registered. This project started in August 2012 and the system became operational in July 2014. The project has participated in multiple student competition and won the second prize ($3,500) in one competition. The project received three grants from the U.S. National Science Foundation for (1) international collaboration, (2) entrepreneurship, and (3) big data analytics. More than a dozen research papers have been published, including one best paper award, and one article featured on the cover of an IEEE magazine. The lead investigator has been invited as a keynote speaker in an international conference and as a panelist in two international conferences. Three students have registered a company to commercialize relevant technologies inspired by this project (but not the project itself). This paper describes the project in detail and shares experiences on many crucial factors in creating a successful cross-cohort research project.

Project Description

Since the beginning of the 21st century, technologies have become mature collecting large quantities of data from many distributed sources. "Big Data" have been used in a wide range of studies, such as customer behavior[10-11]. Among all types of data, visual data are, arguably, the "biggest". Commercial digital cameras have resolutions more than 20 million pixels. A HD (high definition) video camera costs less than $100 USD and can continuously stream data at 30 frames per second. Millions of network cameras are deployed each year. Some of them are used for surveillance purposes (also called closed-circuit television or CCTV) and only authorized personnel can access the data. Meanwhile, many network cameras are connected to the Internet without password protection. Examples include traffic cameras[12], air quality cameras[13], and meteorology[14]. The data provides great resources for understanding the world (both natural and manmade)[15]. We are not aware of any project that can systematically exploit the visual data on the Internet. The closest project we know is AMOS[16] but it has low frame rates and does not provide real-time data. This VIP project started in August 2012 by a faculty member and two undergraduate students (both are doctoral students now). The team grows gradually and regularly has more than 10 students per semester. More than 50 students have contributed to this project; many of them participate multiple semesters and several undergraduate students have been in the team for more than two years.

The architecture of this project includes four major components: a web server, a camera database, a user database, and a resource manager[17-18]. The web interface uses Django and the database uses MySQL. The entire system is written in Python, using OpenCV for data analysis. The computing engines are cloud instances; Amazon and Microsoft support this project by providing research credits. The system can access more than 80,000 network cameras. The system provides an application programming interface (API) through which users can write their programs analyzing the real-time visual data. The largest experiment so far detects motion in 16,000 network cameras at one frame every five seconds for 24 hours; approximately 200 million images are processed. Interested readers may register as users at the project's web site: Continuous Analysis of Many Cameras, CAM^2, https://cam2.ecn.purdue.edu/.

Project Management

Research is not a requirement for the undergraduate degree in our department. Thus, recruiting talent is one of the most important factors for the success of a project. The lead investigator teaches a 200-level (sophomore) programming class and this class is a channel to recruit students. When a student expresses interest in this project, the investigator explains the vision of the project: it is important for a student to understand the purpose of a project because each student has only limited time outside coursework. This project aims to build a tool available to researchers observing real-time phenomena worldwide. Building something for other people is appealing to many students, as evident in service learning[4-5]. As more students have participated in this project, internal referral becomes another important channel for recruiting. Some team members encourage their friends to join the team and work together. In typical classroom settings, all students solve the same problems. This is, however, different from the "real world" where different team members solve different problems in the project. This project provides an opportunity for students to understand how "real world" projects operate. Sometimes, a member is unable to accomplish assigned tasks and the other members have to develop a contingency plan. All these factors make this project attractive to students because they can acquire the skills managing projects.

Students also learn software management skills that are commonly used in the software industry. As this project has many team members (and international collaborators), students have to be proficient in using version control. This project adopts `git` because it is distributed, allowing members to work on different parts of the project simultaneously. The commit-push process allows a student to commit changes without affecting the other students, as long as the changes are not pushed. Few students have experience with version control, due to the nature of class projects (individual assignments and short durations). Some students use version control primarily for backing up code, not as a way to collaborate. Another essential skill that students need to acquire is debugging. Many large software projects contain defects that need to be discovered and corrected. Individual developers should use unit testing to verify that their code performs to specifications and is free from defects. Learning how to test for and discover defects (especially difficult "heisenbugs" that can elude testing and debugging) is a skill that is best learned by doing. The next skill needed is creating a production quality software testing and release environment. A sizable effort is required for creating test scripts, running a comprehensive suite of regression tests on a frequent basis. The project today includes the development and use of test scripts along with limited regression tests. As this paper is written, significant efforts are being taken for creating testing scripts that allow *continuous integration* (CI). CI is a new concept for almost all students. Students appreciate the opportunities learning the best practice adopted in software industry.

Within the broad area of working within a team, many non-technical skills are needed. The first is understanding the dynamics of team formation and how to work within a large team of developers. Familiarity with how teams form and how to work best within a team, including the use of communication tools, is a critical skill that will help students transition into working in industry. One major element is time management, working on tasks within time constraints. Learning how to estimate task durations, effectively communicate task status, and manage working time and risk factors to stay within a project schedule time constraint is an essential skill. Entrepreneurship is encouraged in today's high education. Three students recognize the market opportunities of analyzing large amounts of video streams and plan to start a technology company when they graduate in May 2016.

Factors for a Successful Cross-Cohort Research Project

Creating a successful research project that includes many undergraduate students requires careful design. The ambitious goal of creating a system using global network cameras is an intellectual challenge. Recruiting students from a sophomore class is a major factor contributing to the success of the project, since these students have the option of staying in the project for multiple semesters. These students can provide the continuity of the project. Meanwhile, these students have limited knowledge, skills, and experience and thus training is essential. Returning students are explicitly assigned the task of training new members. Returning members also have opportunities to be promoted to lead subteams or the entire team. In fact, the leadership succession plan is determined before each semester so that the new leaders can start planning and preparing training materials. Current team members interview new members, evaluate their expertise, and determine the most appropriate tasks for each new member. The experience of being interviewers greatly improves students' understanding how companies select candidates. Undergraduate students have much higher turnover than graduate students because undergraduate students are still searching for their career aspiration. Disruptions may occur for many reasons: internship, co-op, study abroad. Thus, succession plans help the team, as well as every student, plan for their future semesters. This project requires many different types of talent: software development, leadership, communication skill, entrepreneurship, etc. This diversity of talents has proven attractive to many students.

Conclusion

This paper describes a cross-cohort research team, the goal of the project, recruitment and management of the team. Such a project can provide learning experience unavailable in typical classroom settings. Many factors contribute to the success of such a project, including management structure, succession plans for leaders, efforts on training new members, using software for management and collaboration, distinction between classroom learning and solving research problems, etc. Even though this project is still at its early stage, it has been able to attract many students. We expect this new model of learning to complement existing classroom learning and help students succeed in workplace.

Acknowledgments

We want to thank the AWS Cloud Credits for Research and the Microsoft Azure for Research Program. This project is supported in part by NSF ACI-1535108, CNS-0958487, and CNS-0855098. Any opinions, findings, and conclusions or recommendations in this materials are those of the authors and do not necessarily reflect the views of the sponsors.

References

1. Nanette Brown, Yuanfang Cai, Yuepu Guo, Rick Kazman, Miryung Kim, Philippe Kruchten, Erin Lim, Alan MacCormack, Robert Nord, Ipek Ozkaya, Raghvinder Sangwan, Carolyn Seaman, Kevin Sullivan, and Nico Zazworka. 2010. Managing technical debt in software-reliant systems. In Proceedings of the FSE/SDP workshop on Future of software engineering research, 47-52. DOI=10.1145/1882362.1882373 http://doi.acm.org/10.1145/1882362.1882373
2. Yuepu Guo and Carolyn Seaman. 2011. A portfolio approach to technical debt management. In Workshop on Managing Technical Debt, 31-34. DOI=http://dx.doi.org/10.1145/1985362.1985370
3. Lim, E.; Taksande, N.; Seaman, C., 2012. "A Balancing Act: What Software Practitioners Have to Say about Technical Debt," in Software, IEEE , vol.29, no.6, pp.22-27, Nov.-Dec. doi: 10.1109/MS.2012.130
4. Coyle, E. J., Jamieson, L. H., Oakes, W. C. 2005. “EPICS: Engineering Projects in Community Service”, International Journal of Engineering Education, 21(1), 139–150.
5. Zoltowski, C. B., & Oakes, W. C. 2014. “Learning by Doing: Reflections of the EPICS Program” [Special Issue]. International Journal of Service Learning in Engineering, Humanitarian Engineering, and Social Entrepreneurship, 9(3), 1–32.
6. Coyle, E., & Krueger, J., & Allebach, J. 2006, “The Vertically Integrated Projects (VIP) Program in ECE at Purdue: Fully Integrating Undergraduate Education and Graduate Research,” In Proceedings of the 2006 American Society for Engineering Education Annual Conference & Exposition.
7. Carolyn Ash Merkel, 2003. “Undergraduate Research at the Research Universities”, New Directions for Teaching and Learning, Issue 93, Wiley Subscription Services, Inc., http://dx.doi.org/10.1002/tl.87
8. Huff, J. L., Zoltowski, C. B. and Oakes, W. C. 2016. “Preparing Engineers for the Workplace through Service Learning: Perceptions of EPICS Alumni,” Journal of Engineering Education, 105: 43–69. doi: 10.1002/jee.20107
9. E.J. Coyle, J.V. Krogmeier, R.T. Abler, A. Johnson, S. Marshall and B.E. Gilchrist, 2014. “The Vertically-Integrated Projects (VIP) Program – Leveraging Faculty Research Interests to Transform Undergraduate STEM Education,” Proceedings of the Transforming Institutions: 21st Century Undergraduate STEM Education Conference,

10. Gediminas Adomavicius, Alexander Tuzhilin, 2001. "Using Data Mining Methods to Build Customer Profiles", Computer, vol.34, no. 2, pp. 74-82, February, doi:10.1109/2.901170
11. Xindong Wu; Xingquan Zhu; Gong-Qing Wu; Wei Ding, 2014. "Data mining with big data," in Knowledge and Data Engineering, IEEE Transactions on , vol.26, no.1, pp.97-107, Jan. doi: 10.1109/TKDE.2013.109
12. New York City Department of Transportation, http://nyctmc.org/
13. National Park Service, Air Quality Web Cameras, http://www.nature.nps.gov/air/webcams/
14. NOAA Great Lakes Environmental Research Laboratory, http://www.glerl.noaa.gov/metdata/
15. Wei-Tsung Su, Yung-Hsiang Lu, and Ahmed Kaseb, 2015. "Harvest the Information from Multimedia Big Data in Global Camera Networks", IEEE International Conference on Multimedia Big Data
16. Nathan Jacobs, Nathaniel Roman, Robert Pless, 2007. "Consistent Temporal Variations in Many Outdoor Scenes", In IEEE Conference on Computer Vision and Pattern Recognition
17. Wenyi Chen, Anup Mohan, Yung-Hsiang Lu, Thomas Hacker, Wei Tsang Ooi, Edward J Delp, 2015 "Analysis of Large-Scale Distributed Cameras Using Cloud", IEEE Cloud Computing, September/October .
18. Wenyi Chen, Yung-Hsiang Lu and Thomas Hacker, 2015 "Adaptive Cloud Resource Allocation for Analysing Many Video Streams", IEEE International Conference on Cloud Computing Technology and Science (CloudCom)