Analysis of Current and Future Computer Science Needs via Advertised Faculty Searches for 2016

Craig E. Wills

Professor and Department Head
Computer Science Department
Worcester Polytechnic Institute
Worcester, MA 01609

WPI-CS-TR-15-03

November 2015

Abstract

This work uses the same methodology as work from last year to study where Computer Science departments are choosing to invest faculty positions by examining data obtained from advertised faculty searches for the current hiring season. While the number and areas for faculty searches does not necessarily translate into the same for faculty hires, we believe that they provide insight into current and future needs within the discipline.

We analyzed ads from 267 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. There is a 20% one-year increase in the number of institutions searching for tenure-track faculty in Computer Science and a 26% increase in the number of positions being searched for. In particular, the number of positions seeking to be filled by top-100 PhD institutions has increased by 29% in one year and the number of positions for undergraduate-only institutions has increased by 46%.

In terms of specific areas, we found that the area clusters of Security, Big Data and Systems/Networking continue to be the areas of greatest investment. We also found that 30-60% of all hires are for areas that are, or may be, interdisciplinary in nature.

Differences are also seen when analyzing results based on the type of institution. Positions related to Security have the highest percentages for top-100 PhD and MS institutions. Big Data is of most interest for other PhD institutions, while Systems/Networking is at the top for BS institutions. Finally, the abundance of potentially interdisciplinary areas is most pronounced for graduate institutions with 30-60% of all positions devoted to these areas.

1 Introduction

The wealth of faculty searches in Computer Science during this hiring season for positions starting in the Fall of 2016 again affords the opportunity to study areas of Computer Science where departments are choosing to invest in new faculty hires. This report details results in a similar manner as a study of faculty hiring ads in Computer Science done one year ago [1]

The primary focus of this work is to study where departments specifically, and the discipline more generally, are choosing to invest precious tenure-track faculty positions. It is an opportunity to understand where Computer Science departments think they are in terms of current needs as well as where they think they are going.

With this focus, there are a number of caveats to our study:

1. Our study is not exhaustive in that it does not necessarily take into account all searches currently underway for this hiring season. We describe the methodology used to discover ads, but ads may have been missed or may not have been placed in the timeframe of our study.

2. While our study focuses on preferred areas for faculty applicants, not all ads identify such preferred areas. These searches are accounted for in the data, but are not considered when analyzing particular areas of interest.

3. Our study analyzes searches and not hires. The number and areas of actual faculty hires may not match what is being searched for.

2 Methodology

We used two primary sources for obtaining ads for Computer Science faculty positions: the Computer Research Association (CRA) Job postings[1] and the Association for Computing Machinery (ACM) list of jobs[2]. We considered ads posted on these venues between August, 2015 and mid-November 2015, which is the same timeframe used in our previous study. In addition, we augmented these two sources with positions posted on the SIGCSE mailing list, which often includes ads for more teaching-focused institutions.

Only ads for tenured and tenure-track positions by departments containing Computer Science were considered. We did not consider non-tenure-track positions such as lecturers, instructors or researchers. We also did not consider searches for department head positions.

For each ad we coded the institution name and the number of positions being searched for. If the ad included specific areas of interest then these were coded as well.

3 Results

Using this methodology our resulting dataset contains information for faculty searches from 267 institutions (249 are U.S. based). 216 (81%) of these institutions indicate a specific number of

[1] http://cra.org/ads/
[2] http://jobs.acm.org/c/search_results.cfm?site_id=1603

positions being searched for with the remaining searches using non-specific phrases such as "multiple positions," "several positions" or just "positions" to indicate the number. As comparison, our previous-year study [1] found searches for 223 institutions (212 U.S. based) with 182 (82%) of these institutions indicating a specific number of positions being searched for. These numbers show a 20% increase over one year in the number of institutions searching for tenure-track faculty in Computer Science.

Similarly not all ads listed specific or preferred areas of interest. 197 (74%) of the 267 institutions listed specific areas, which is a comparable percentage as last year. In studying particular areas of interest, we only considered the ads from these institutions for our analysis.

3.1 Results by Area

In the initial step of our study, we determined the number of times that a specific area was mentioned in an ad. Thus an ad for a single faculty position with preferred interest in the areas of HCI, Security, Machine Learning and Robotics would count one "mention" for each of these four areas. Another institution looking to focus three positions in the area of Security would be one mention for Security. A total of 739 specific areas are mentioned in ads (versus 488 last year). Figure 1 shows the percentage of mentions for each area in this year's data.

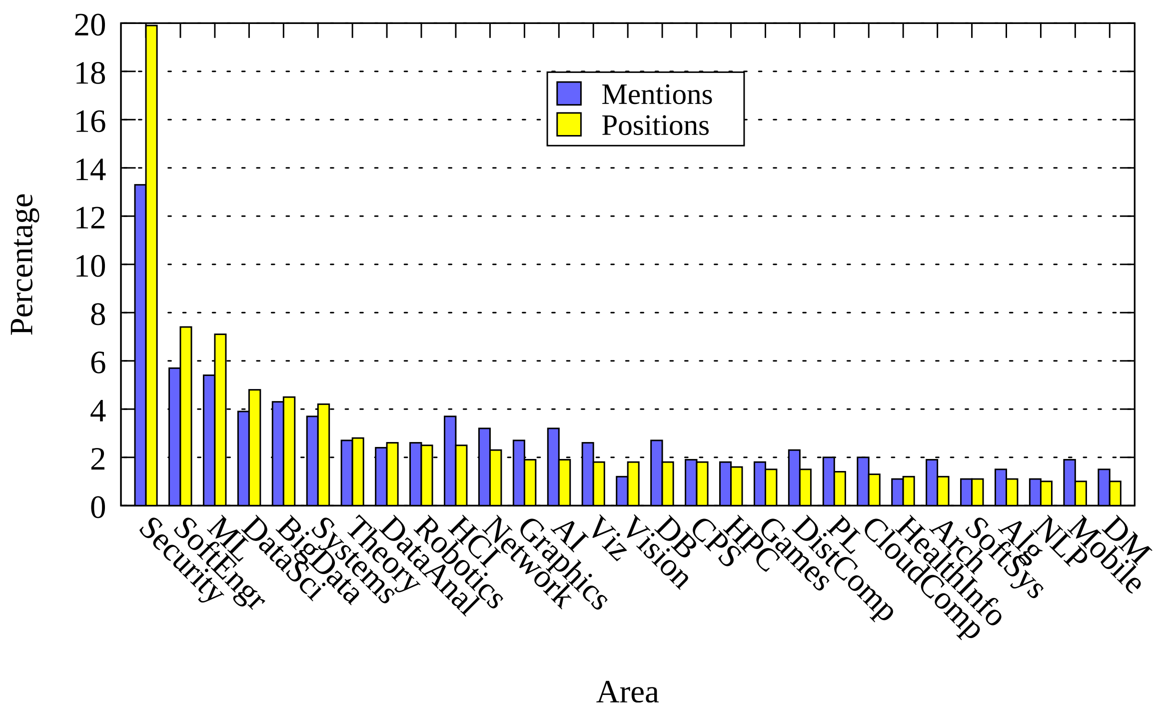

Figure 1: Area Percentage by Mentions and by Positions

While mentioned areas are one metric, another approach is to consider a faculty search as a "vote" for an area of current and future need. Using this approach a single position with four areas of interest would be investing 0.25 positions for each area, while three positions focused in a single area would invest 3.0 positions in that single area.

The problem with weighting areas based on the number of positions is that not all ads list a specific number of positions. For ads indicating multiple positions there are at least two, but otherwise the number is not known. To simplify our analysis we mapped each of these "multiple positions" searches to a fixed number of positions. We experimented with fixed values between two and four positions without significant differences in the overall results. All results shown in this report use a fixed value of three for multiple-position searches resulting in a total of 507 "positions" being searched for by the 267 institutions with 380 (75%) of the positions indicating preferences for specific areas. Last year there were 401 such positions with 310 (77%) indicating preferences for specific areas. The change from 401 to 507 positions being searched for represents a 26% increase over one year.

Figure 1 also shows the percentage of positions for areas with at least one percent for either mentions or positions. They are shown in rank order based on the number of positions. The results show that the area of Security, which includes Privacy, accounts for the highest percentage of both mentions and positions, although it accounts for relatively more positions. Big Data is the area with the second most number of positions, although Software Engineering has the second most number of mentions. These relative positions for areas are the same as last year. Apart from these two areas, the percentages for each metric are comparable for most of the remaining areas, although the percentage of mentions is a bit higher for areas such as HCI, Mobile Systems and Graphics.

3.2 Results by Area Cluster

Figure 1 does not show other mentioned areas that appear less frequently in ads. In order to account for these less-frequent areas and to combine similar areas, such as Big Data and Data Science, we grouped areas into clusters. These area clusters and the set of areas constituting the cluster are shown in Table 1. Areas with a small number of mentions and not clearly fitting into a cluster are included in two other clusters—one with areas in traditional Computer Science (OtherCS) and one with areas more interdisciplinary in nature (OtherInter). These are the same clusters used in our previous study [1].

Figure 2 shows the same results as Figure 1 using the clusters from Table 1 rather than the areas directly. It shows that the Security area cluster has the highest percentage of positions while Big Data has the highest percentage of mentions. The Systems/Networking area cluster is in the top three for each metric. These three area clusters were also the top three by percentage last year, although Security has overtaken Big Data as the top cluster based on positions. The remaining area clusters tend to have more comparable percentages between the two metrics save for Software Engineering, HCI and Mobile Systems.

3.3 Results Comparison with Previous Year

Figure 3 shows a more complete comparison of clustered area results based on percentage of positions for the previous year searches of 2015 with the current year searches for 2016. Clustered areas are ordered based on 2016 percentages. The results show that an even higher percentage of positions are being targeted for Security hires with a relatively lower percentage of positions for the Big Data and Systems/Networking clusters. There is a relative increase in the number of positions for AI with a relatively smaller percentage for Software Engineering. For the remaining

Table 1: Areas Grouped in Each Area Cluster

Area Cluster	Constituent Areas
AI	AI, Data Mining, Machine Learning
Arch	Architecture
BigData	Big Data, Data Science, Data Analytics, Data Computation/Systems, Information Analysis, Visualization, Visual Computing, Knowledge Representation
Bioinfo	Bioinformatics
Compiler/PL	Compilers, Programming Languages, Object-Oriented Languages
CompSci	Computational Biology, Computational Life Science, Computational Science
DB	Database, Data Management
Games	Games, Interactive Media, Digital Media
HCI	HCI, Human Computing, Interactive Computing, Virtual Reality
ImageSci	Graphics, Image Processing, Pattern Recognition, Vision
Mobile	Mobile Systems, Ubiquitous/Pervasive Computing
Robotics/CPS	Autonomous Systems, Cyber-Physical Systems, Embedded Systems,Internet of Things, Reconfigurable Systems, Robotics, Sensors
Security	Forensics, Privacy, Security
SoftEngr	Software Assurance, Software Design, Software Engineering, Software Systems
Sys/Net	Cloud Computing, Distributed Computing, High Performance Computing, Experimental Systems, Networking, Network Science, Operating Systems, Parallel Computing, Systems
Theory/Alg	Algorithms, Theory, Formal Methods
OtherCS	Applied Areas, CS Education, Informatics, Information Technology, Natural Language Processing, Software, Social Networking, System Design, System Verification, Web Technologies
OtherInter	Bioscience, Biomedical, Business Intelligence, Cognitive Model, Cognitive Systems, Communications, Energy Awareness, Geographic Information Systems, Health Computing, Health Information Systems, Interdisciplinary, Learning Science, Neuroscience, Social Computing, Urban Informatics

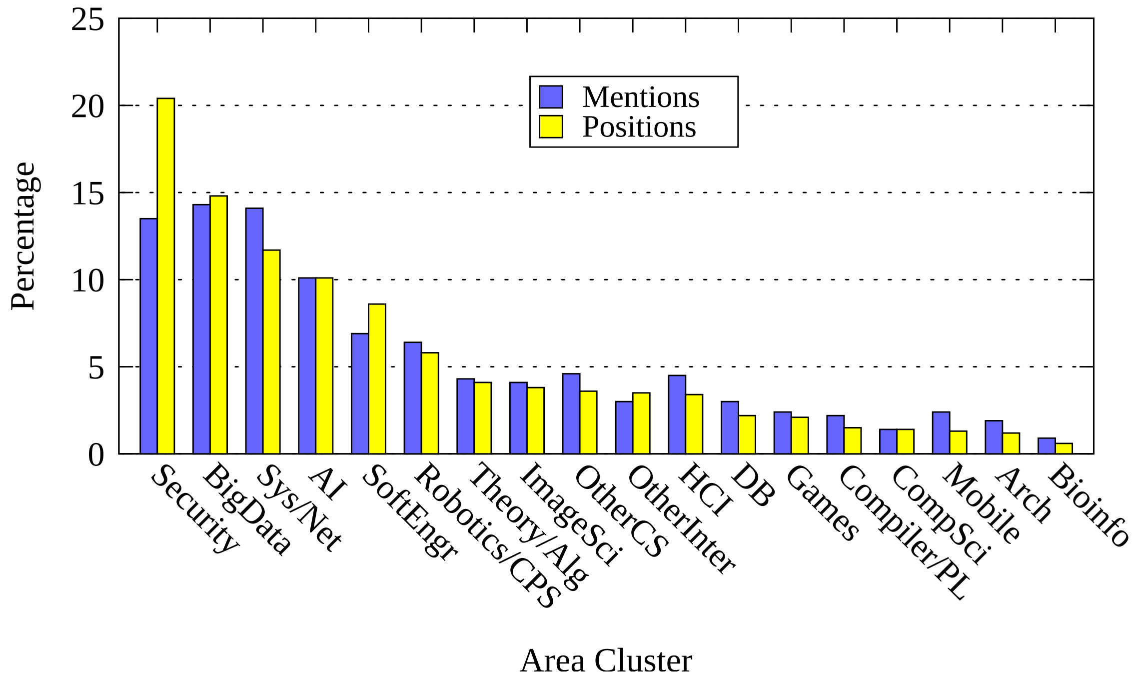

Figure 2: Clustered Area Percentage by Mentions and by Positions

area clusters, there are relatively larger year-over-year increases in Theory/Algorithms and Image Sciences with relatively larger year-over-year decreases in Databases and Bioinformatics.

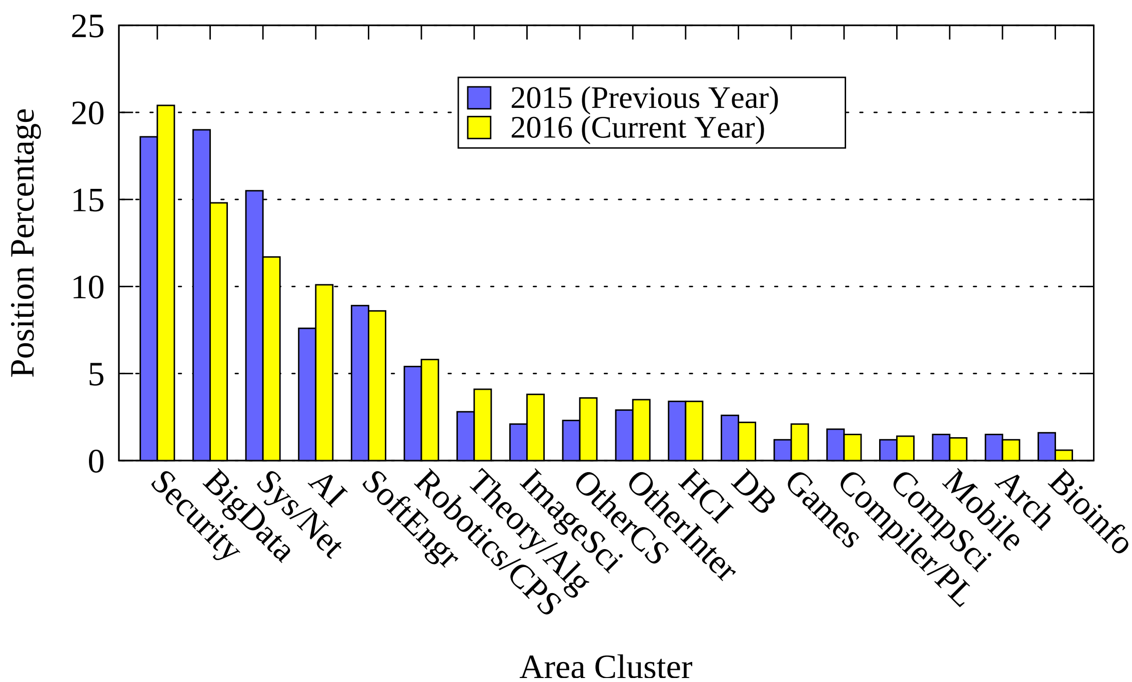

Figure 3: Clustered Area Percentage by Positions for 2016 in Comparison with 2015

3.4 Results for Interdisciplinary Area Clusters

Another question we again examined is how the interdisciplinary nature of Computer Science is affecting hiring. Specific clusters in Table 1 that are more interdisciplinary include the Big Data, Robotics/CPS, Bioinformatics, Games, Computational Science, and Other Interdisciplinary clusters. Combining the results for these clusters from Figure 2 shows that 28% of the mentions and 28% of the positions are for these more interdisciplinary areas. Moreover, other clusters such as AI and Security either support interdisciplinary work or may include work with other disciplines. Including these two clusters, which have some amount of interdisciplinary nature, results in up to 52% of the mentions and 59% of the positions being interdisciplinary in nature. These numbers are comparable to previous year results.

4 Results By Type of Institution

We repeated our analysis based on the type of the program at each institution. For example, undergraduate-only programs may not have the same needs as PhD programs. For this portion of the study we augmented our dataset to include the highest degree offered by each program—BS, MS or PhD. Our dataset includes 146 PhD institutions—up from 122 last year. In order to study faculty investments at the most prominent U.S. programs, we further subdivided this group by using the U.S. News Rankings of the 100 Best Graduate schools[3]. This "PhD100" list account for 85 (vs. 70 last year) institutions in our dataset. The remaining PhD programs, including the 18 non-U.S. based, are denoted as "PhDOther". Table 2 shows summary results based on the four institution types.

Table 2: Summary of Position Searches by Institution Type

Institution Type	Number of Institutions	Advertised Number of Positions 1	2	3+	Multiple	Total Positions	%Specific Area
PhD100	85	19 (22%)	20 (24%)	12 (14%)	34 (40%)	216	73%
PhDOther	61	26 (43%)	16 (26%)	7 (11%)	12 (20%)	125	90%
MS	46	28 (61%)	9 (20%)	5 (11%)	4 (9%)	80	78%
BS	75	66 (88%)	8 (11%)	0 (0%)	1 (1%)	86	56%
All	267	139 (53%)	53 (21%)	24 (7%)	51 (18%)	507	75%

Table 2 reveals differences between the different types of institutions. Ads for 88% of the BS institutions are for a single position while 40% of the ads for PhD100 institutions are for multiple positions. The overall percentages are generally comparable as last year.

As shown, the distributions translate into a total number of 216 positions for PhD100 institutions, which is a 29% increase from last year. The position counts and relative change from last year for the remaining types are 125 positions (18% increase) for PhDOther, 80 positions (16% increase) for MS and 86 positions (46% increase) for BS institutions.

[3]http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-science-schools/computer-science-rankings

The last column of Table 2 shows that only 56% of positions from BS institutions identify specific areas of interest while 90% of PhDOther institutions do so. The percentage for PhD100 institutions is 73%, which is a decrease from 78% last year.

In order to understand differences on areas of interest between different types of institutions for 2016 searches, we use the position metric results and area clusters shown for all institutions in Figure 2. Figure 4 shows these same data (in the same rank order as Figure 2) grouped by the four types of institutions.

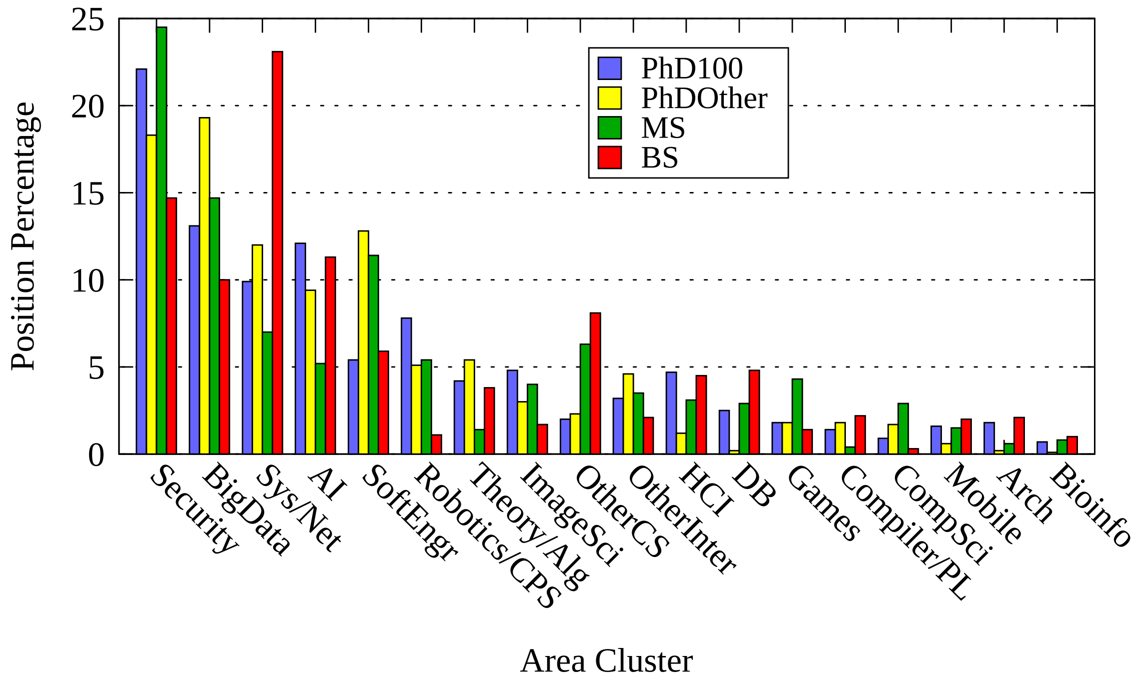

Figure 4: Area Cluster Percentage by Institution Type

The figure shows a number of interesting results. Positions related to Security have the highest percentages for PhD100 and MS institutions. Big Data is of most interest for PhDOther institutions, while Systems/Networking is at the top for BS institutions. AI is more in demand for PhD100 and BS institutions while Software Engineering is more in demand for PhDOther and MS institutions. The Robotics/CPS cluster is relatively much more in demand by the graduate institutions. Compared to last year many of the results are the same, although Security overtook Big Data and Systems/Networking as the most-sought-after area cluster for PhD100 and MS institutions. Systems/Networking overtook Security as the most popular area cluster for BS institutions.

Finally, Figure 5 shows the percentage of positions devoted to areas that are more and some amount of interdisciplinary in nature. The results show that the impact of interdisciplinary areas is even more pronounced for graduate institutions with at least 30% and up to 60% of all positions devoted to these areas. These percentages are comparable to last year for PhD institutions and significantly higher for MS institutions. BS institutions show at most 40% of the positions devoted to these areas, which is a bit lower than last year.

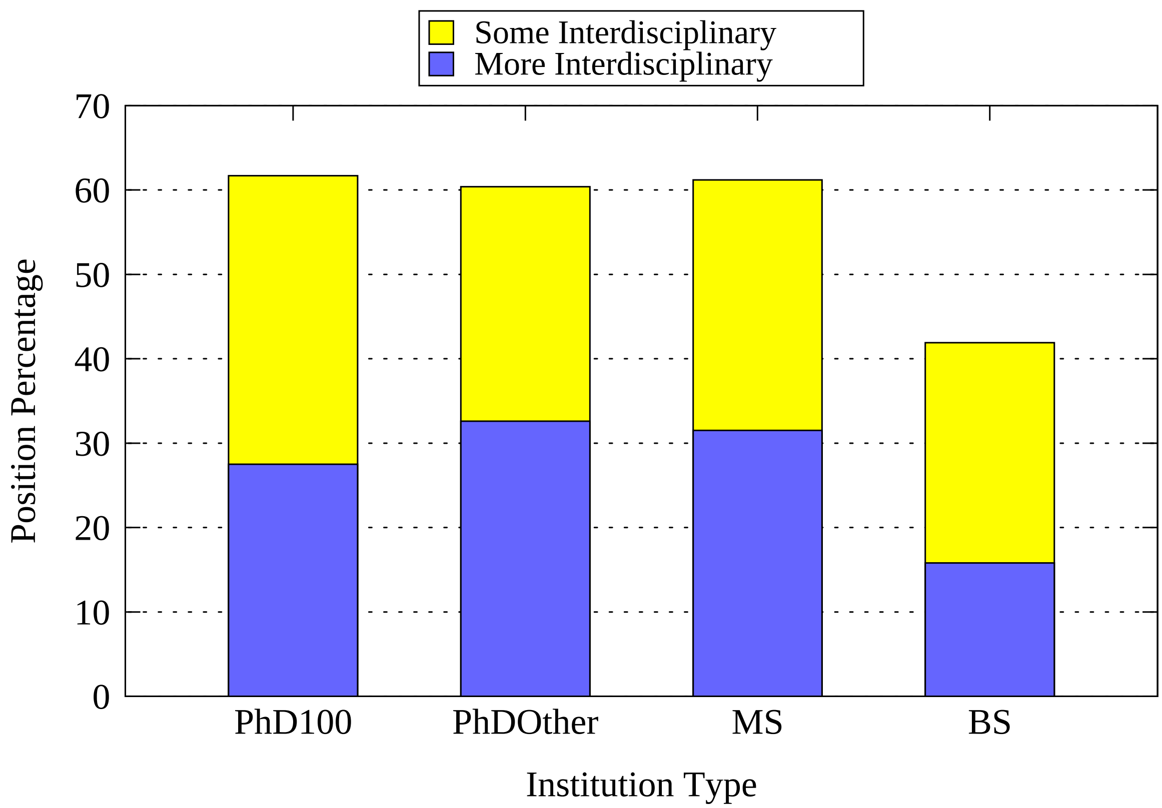

Figure 5: Investment in Interdisciplinary Area Clusters by Institution Type

5 Summary

This work uses the same methodology as work from last year to study where Computer Science departments are choosing to invest faculty positions by examining data obtained from advertised faculty searches for the current hiring season. While the number and areas for faculty searches does not necessarily translate into the same for faculty hires, we believe that they provide insight into current and future needs within the discipline.

We analyzed ads from 267 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. There is a 20% one-year increase in the number of institutions searching for tenure-track faculty in Computer Science and a 26% increase in the number of positions being searched for. In particular, the number of positions seeking to be filled by top-100 PhD institutions has increased by 29% in one year and the number of positions for undergraduate-only institutions has increased by 46%.

In terms of specific areas, we found that the area clusters of Security, Big Data and Systems/Networking continue to be the areas of greatest investment. We also found that 30-60% of all hires are for areas that are, or may be, interdisciplinary in nature.

Differences are also seen when analyzing results based on the type of institution. Positions related to Security have the highest percentages for top-100 PhD and MS institutions. Big Data is of most interest for other PhD institutions, while Systems/Networking is at the top for BS institutions. Finally, the abundance of potentially interdisciplinary areas is most pronounced for graduate institutions with 30-60% of all positions devoted to these areas.

References

[1] Craig E. Wills. Analysis of current and future computer science needs via advertised faculty searches. *Computing Research News*, 27(1), January 2015. Full report at `http://web.cs.wpi.edu/~cew/papers/CSareas15.pdf`.