Analysis of Current and Future Computer Science Needs via Advertised Faculty Searches for 2018

Craig E. Wills

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

WPI-CS-TR-17-04

November 2017

Abstract

This work uses the same methodology as work from previous years 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 of 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 406 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. There is a 17% one-year (and 52% two-year and 82% three-year) increase in the number of institutions searching for tenure-track faculty in Computer Science and a 21% one-year (and 64% two-year and 107% three-year) increase in the number of positions being searched for. The number of institutions searching and positions seeking to be filled has increased the most for BS institutions.

In terms of specific areas, we found that the clustered areas of Security, Data Science and AI/Data Mining/Machine Learning are the areas of greatest investment. Aggregating the Data Science, AI/DM/ML and Databases clusters results in close to one-third of all hires sought in these data-oriented areas. We again 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, MS and BS institutions. Data Science is of most interest for other PhD institutions. 35% of positions for PhD institutions are in data-oriented areas. Finally, the abundance of potentially interdisciplinary areas is most pronounced for PhD institutions with 32-65% 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 2018 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 in previous years [1, 2, 3]. The longitudinal aspect also allows insight into the number of positions and the areas being sought over time.

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

As in past years, we used three primary sources for obtaining ads for Computer Science faculty positions: the Computer Research Association (CRA) Job postings[1] the Association for Computing Machinery (ACM) list of jobs[2], and the Chronicle of Higher Education Vitae site[3]. We again augmented these sources with positions posted on the SIGCSE mailing list, which often includes ads for more undergraduate-focused institutions. We considered ads posted on these venues between August 2017 and mid-November 2017, which is the same timeframe used in our previous studies.

Only ads for tenured and tenure-track positions by departments containing Computer Science or closely-related programs were considered. We did not consider non-tenure-track positions such as lecturers, instructors or researchers and we only considered institutions awarding at least a BS degree. Searches for Deans or department chair positions were noted, but not considered because they do not reveal information regarding areas. Similarly, searches for other departments and programs with interest in faculty with Computer Science background were noted, but also not considered.

[1] https://cra.org/ads/
[2] http://jobs.acm.org/jobs/search
[3] https://chroniclevitae.com/job_search/new

3 Results

3.1 Institutions and Positions

Using this methodology our resulting dataset contains information for faculty searches from 406 institutions (356 are U.S. based). 318 (78%) of these institutions indicate a specific number of 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 [3] found searches for 347 institutions (313 U.S. based) with 78% of these institutions indicating a specific number of positions being searched for.

The left-side of Figure 1 shows four-year results for the number of institutions searching. It shows a 17% increase over the past year, a 52% increase over the past two years, and a 82% increase over the past three years.

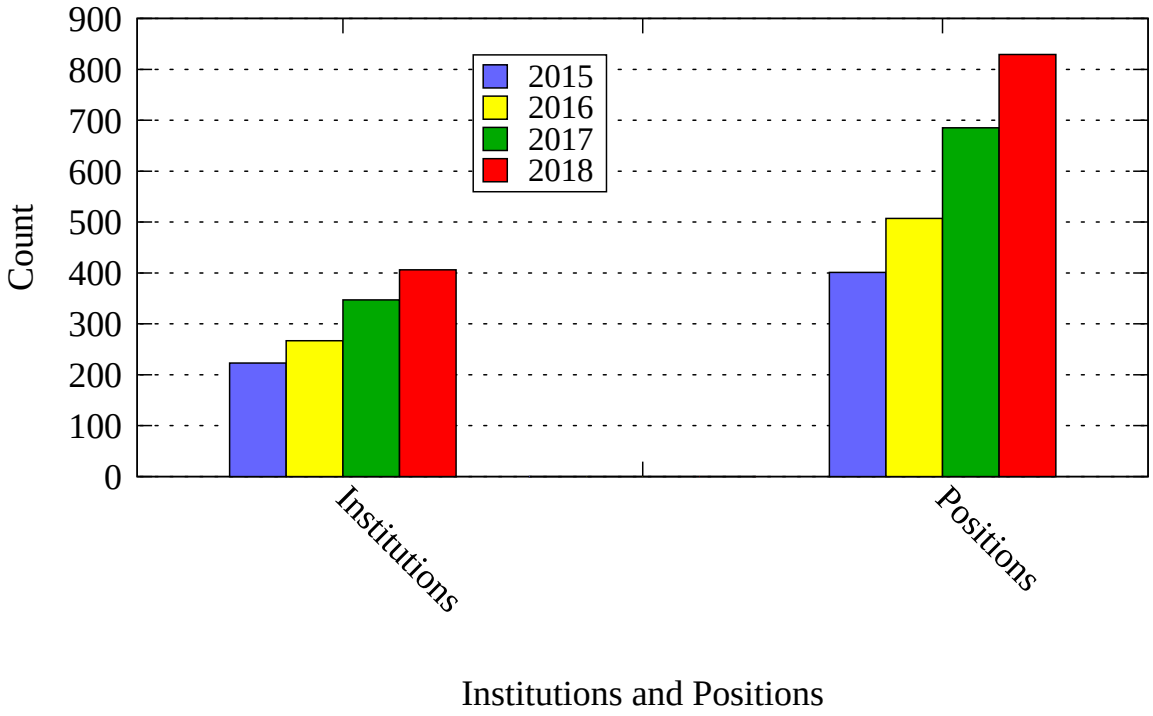

Figure 1: Longitudinal Results for Number of Institutions Searching and Total Number of Positions Being Searched For

In terms of the total number of positions, in the past we experimented with treating such "Multiple Position" searches as meaning two, three or four positions and settled on a value of three. In related work [4] where we surveyed institutions on their hiring outcomes we found those seeking multiple positions responded seeking a median of 3 and mean of 3.5 positions. The right-side of Figure 1 shows the total number of positions searched for using a value of three for "multiple positions" for a total of 829 positions. This value represents a 21% one-year, a 64% two-year, and a 107% three-year increase in the number of positions being searched for.

Finally, in terms of institutions and positions we did encounter additional ads for Computer Sci-

entists that were noted, but not considered in our analysis. We found 27 Dean and Chair searches (up from 20 last year) as well as 146 faculty searches in other departments. These were primarily in Electrical & Computer Engineering or Information School/Science/Technology departments, although also included Bio/Health and Business departments. We do not represent that this list of additional searches is complete because it based only on ads in the venues we describe in our methodology. Ads appearing in other venues were not considered.

3.2 Results by Topic

In the same way that not all ads list a specific number of positions, it is also the case that not all ads list specific or preferred topics of interest[4]. 310 (76%) of the 406 institutions listed specific topics, which is up from 71% last year. In studying particular topics of interest, we only considered the ads from these institutions for our analysis.

In the initial step of our study, we determined the number of times that a specific topic was mentioned in an ad. Thus an ad for a single faculty position with preferred interest for the topics of HCI, Security, Machine Learning and Robotics would count one "mention" for each of these four topics. Another institution looking to focus three positions for the topic of Security would be one mention for Security. A total of 1289 specific topics are mentioned in ads (versus 1074 last year).

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

The problem with weighting topics based on the number of positions is that not all ads list a specific number of positions. We again use the fixed value of three for multiple-position searches resulting in a total of 829 "positions" being searched for with 629 (75%) of the positions indicating preferences for specific topics. Figure 2 shows the percentage of mentions and positions for topics 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 topic of Security accounts for the highest percentage of both mentions and positions, although it accounts for relatively more positions. Security was also the top topic for both metrics last year. Data Science is again the topic with the second most percentage of positions, but drops to third for percentage of mentions. Machine Learning is the topic with second most percentage of mentions and third most percentage of positions.

3.3 Results Clustered into Areas

Figure 2 does not show topics that appear less frequently in ads nor does it group similar topics, such as Data Science and Big Data. To address these issues, we clustered topics into areas. These clustered areas and the set of topics constituting the area are shown in Table 1. Topics with a small number of mentions and not clearly fitting into a cluster are included in two other clusters—one with topics in traditional Computer Science (OtherCS) and one with topics more interdisciplinary

[4]In a change from previous reports, we use the term "topic" to refer to sub-domains of Computer Science listed in ads and the term "area" to refer to a clustering of topics.

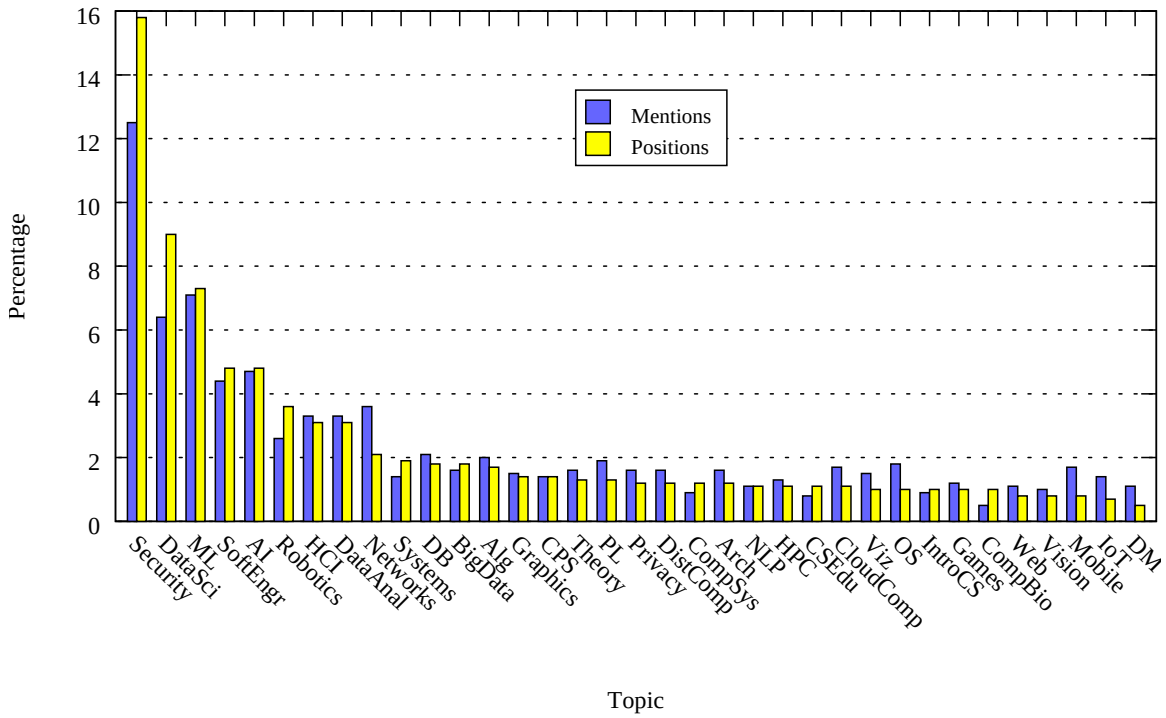

Figure 2: Topic Percentage by Mentions and by Positions

in nature (OtherInter). These are the same clusters used in last year's study [3] and essentially the same as used in previous year studies [1, 2].

Figure 3 shows the same results as Figure 2 using the clusters from Table 1 rather than the topics directly. It shows that the Security clustered area has the highest percentage of mentions and positions. Data Science has the next highest percentage of positions and third highest percentage of mentions. AI/DM/ML has the second highest percentage of mentions and third highest percentage of positions. Compared to last year, the Sys/Net clustered area dropped from third to fourth in percentage of mentions and positions.

The right-most clustered area in Figure 3 was added for this year's results. It is the Data Oriented cluster that further aggregates results for the Data Science, AI/DM/ML and Databases clusters. This aggregated cluster was introduced because of overlap between the three data-oriented clusters while still retaining the three distinct clusters as defined in Table 1. As shown in the figure, this aggregated cluster accounts for 31% of the mentions and 32% of the positions.

3.4 Results Comparison with Previous Years

Figure 4 shows a more complete comparison of clustered area results based on percentage of positions for the previous year searches of 2015, 2016 and 2017 with the current year searches for 2018. Clustered areas are ordered based on 2018 percentages. The results show a bit of a drop for the percentage of positions are being targeted for Security hires and a comparable percentage of positions for the Data Science cluster. The AI/DM/ML cluster jumped significantly with close to 15% of positions while the Systems/Networking cluster is next with a slight drop in percentage.

Table 1: Topics Grouped in Each Clustered Area

Area	Constituent Topics
AI/DM/ML	Artificial Intelligence, Computational Linguistics, Data Mining, Deep Learning, Machine Learning, Natural Language Processing
Arch	Architecture, Hardware
Bioinfo	Bioinformatics
Compiler/PL	Compilers, Programming Languages
CompSci	Biodesign, Biomedical, Computational Biology, Computational Life Science, Computational Neuroscience, Computational Science, Network Science, Numerical Analysis, Scientific Computation
DataSci	Big Data, Data Analytics, Data Science, Visualization, Visual Analysis/Computing
DB	Database, Data Management, Information Retrieval, Information Systems
Games	Animation, Games
HCI	Augmented Reality, Cognitive Science, Disability Technology, HCI, Immersive Systems, Interactive Applications, Virtual Reality
ImageSci	Graphics, Medical Imaging, Pattern Recognition, Vision
Mobile	Human-Centered Computing, Mobile Systems
Robotics/CPS	Autonomous/Vehicular Systems, Cyber-Physical Systems, Embedded Systems, Intelligent Systems, Internet of Things, Robotics, Smart Systems
Security	Cryptography, Forensics, Information Assurance, Privacy, Security, Trusted Computing
SoftEngr	Dependable Software, Software Assurance, Software Design, Software Deveopment, Software Engineering, Software Systems
Sys/Net	Cloud Computing, Computer Systems, Distributed Computing, High Performance Computing, Network/System Administration, Networking, Operating Systems, Parallel Computing, System Analysis, Systems
Theory/Alg	Algorithms, Computational Geometry, Formal Methods, Logic, Theory
OtherCS	CS Education, Data Structures, Information Technology, Internet, Introductory CS, Modeling, Optimization, Quantum Computing, Simulation, Social Computing, Software, Verification, Web Technologies
OtherInter	Applications, Climate Informatics, Computational Engineering, Design Theory, Digital Computation Studies, Digital Health, Digital Libraries, Economic Computing, Ethics, Financial Technology, Interdisciplinary Applications, Journalism Learning Science/Technology, Materials, Operations Research, Statistics, Sustainability

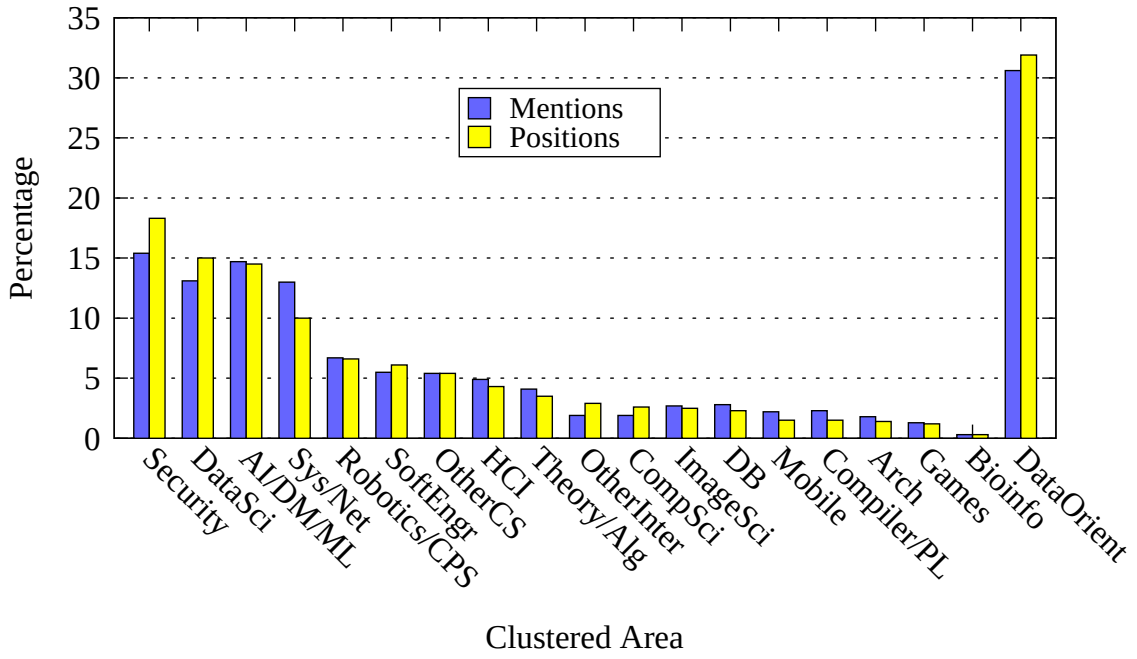

Figure 3: Clustered Area Percentage by Mentions and by Positions

For other topic clusters, Human-Computer Interaction had the largest year-over-year percentage increase while Software Engineering, Mobile Systems and Robotics/CPS had the largest year-over-year decreases. The aggregated Data Oriented cluster increased significantly compared to reconstructed results for previous years. This increase is directly due to the significant percentage increase for the AI/DM/ML cluster.

In terms of longer-term trends, Figure 4 shows with percentage drops for Security, Robotics/CPS and Computational Science, that only Other Computer Science has increased its contributions each year. In contrast, Systems/Networking and Software Engineering have decreased their contributions each of the past three years.

3.5 Results for Interdisciplinary Clustered Areas

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 Data Science, Robotics/CPS, Bioinformatics, Games, Computational Science, and Other Interdisciplinary clusters. Combining the results for these clusters from Figure 3 shows that 29% of the positions are for these more interdisciplinary areas. Moreover, other clusters such as AI/DM/ML 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 62% of the positions being interdisciplinary in nature. The total percentages are a bit higher compared to the previous year results of 59%.

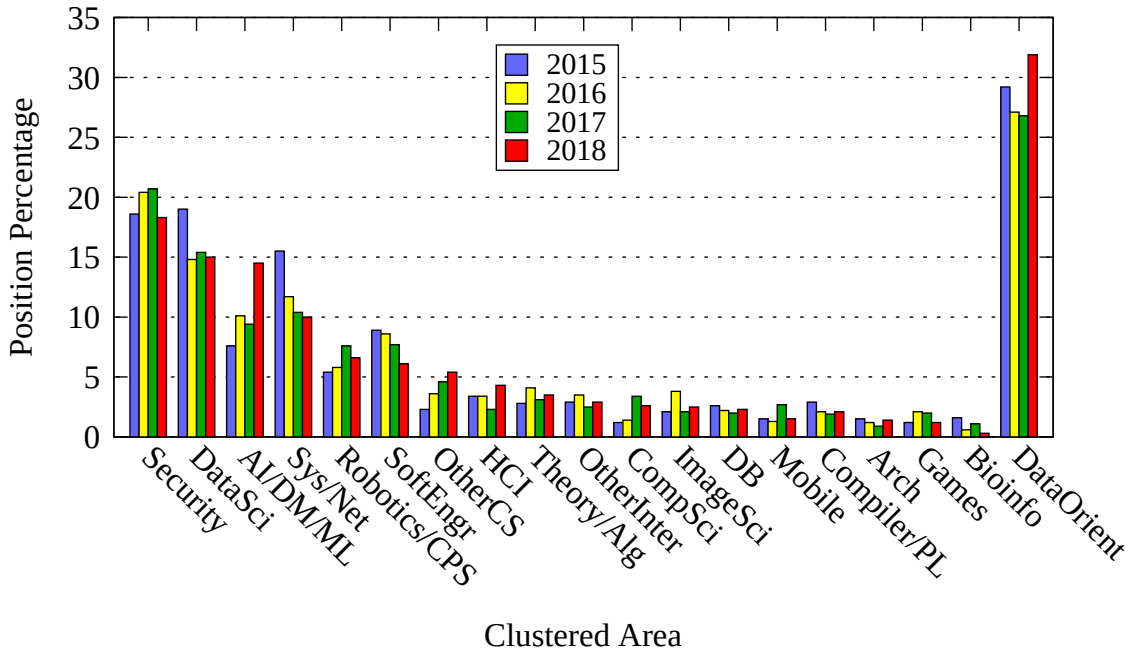

Figure 4: Clustered Area Percentage by Positions for 2018 in Comparison with Previous Years

4 Results By Type of Institution

As means to better understand the results we augmented the dataset to include additional information about each institution. Table 2 shows a breakdown of results based on whether the institution is public, private on non-U.S.-based. The results show that 196 (48%) of the institutions are public and account for 405 (49%) of the total positions. The included non-U.S.-based institutions tend towards seeking higher numbers of positions.

Table 2: Summary of Position Searches by Public, Private, Non-U.S. Institution

Institution Type	Number of Institutions	Advertised Number of Positions 1	2	3+	Multiple	Total Positions	% Positions w/ Specific Topic
Public	196	102 (52%)	34 (17%)	19(10%)	41 (21%)	405	79%
Private	160	100 (62%)	21 (13%)	12(8%)	27 (17%)	295	68%
Non-U.S.	50	12 (24%)	5 (10%)	13(26%)	20 (40%)	129	85%
All	406	214 (53%)	60 (15%)	44(11%)	88 (22%)	829	76%

A more significant distinction is to characterize each institution based on the type of the degree program it offers. 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 178 PhD institutions—up from 165 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[5]. This "PhD100" list accounts for 84 (vs. 86 last year) institutions in our dataset. The remaining PhD programs, including 45 non-U.S. based, are denoted as "PhDOther". Table 3 shows summary results based on the four institution types. The left-side of Figure 5 shows longitudinal results for the number of institutions searching where there is a similar number of PhD100 institutions, but increases for the other types with PhDOther up by 19%, MS by 17% and BS by 29%.

Table 3: Summary of Position Searches by Institution Type Based on Highest Degree Offered

Institution Type	Number of Institutions	Advertised Number of Positions 1	2	3+	Multiple	Total Positions	% Positions w/ Specific Topic
PhD100	84	14 (17%)	8 (10%)	13 (8%)	49 (58%)	294	74%
PhDOther	94	28 (30%)	15 (16%)	20(21%)	31 (33%)	229	88%
MS	83	51 (61%)	18 (22%)	8(10%)	6 (7%)	130	78%
BS	145	121 (83%)	19 (13%)	3 (2%)	2 (1%)	176	62%
All	406	214 (53%)	60 (15%)	44(11%)	88 (22%)	829	76%

Table 3 reveals differences between the different types of institutions. Ads for 83% of the BS institutions are for a single position while 58% 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 294 positions for PhD100 institutions, which is a 9% increase from last year. We note that this number is particularly sensitive to the number of positions assumed for "multiple position" searches as over half of these searches are not specific in the number of positions being sought. The right-side of Figure 5 shows four-year results for the number of positions being searched for by the four types of institutions. The other position counts and relative change from last year are 229 positions (34% increase) for PhDOther, 130 positions (10% increase) for MS and 176 positions (39% increase) for BS institutions.

The last column of Table 3 shows that only 62% of positions from BS institutions identify specific areas of interest while 88% of PhDOther institutions do so with the percentages for the other institution types in between. In order to understand differences on areas of interest between different types of institutions for 2018 searches, we break down the results in Figure 3 based upon the type. Figure 6 shows the results (in the same rank order as Figure 3) grouped by the four types of institutions.

The figure shows a number of interesting results. For the second consecutive year, positions related to Security have the highest percentages for all but PhDOther institutions. Data Science is of most interest for PhDOther institutions and is also of much interest for PhD100 and MS institutions. AI/DM/ML is also of high interest to PhD100 and PhDOther institutions. These results translate into roughly 35% of positions being sought by PhD-granting institutions are for the Data Oriented aggregated cluster, while roughly 25% of positions for BS and MS institutions are in the Data Oriented cluster.

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

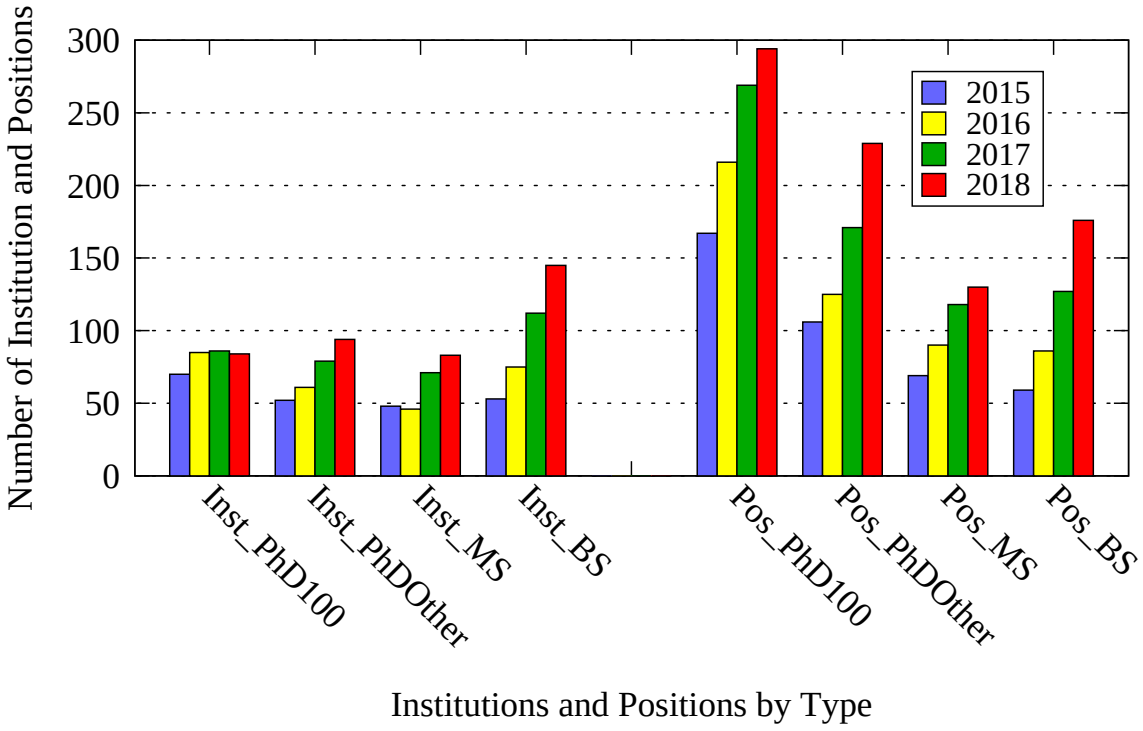

Figure 5: Three-Year Numbers of Institutions Searching and Positions Being Searched For Based on Type of Institution Based on Highest Degree Offered

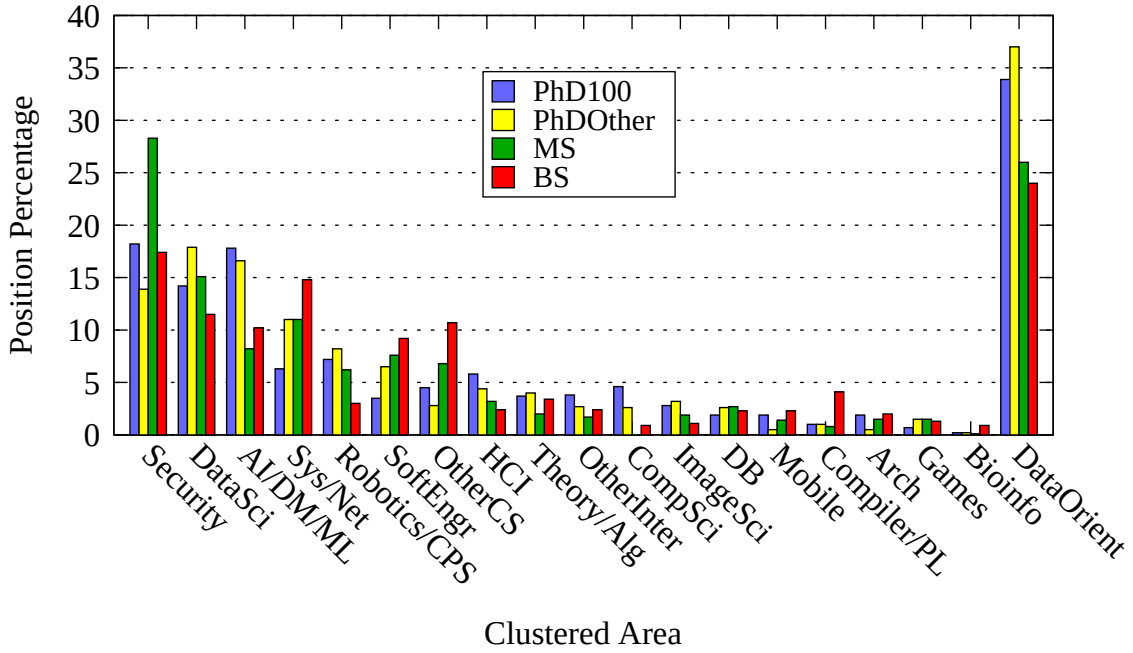

Figure 6: Clustered Area Percentage by Institution Type Based on Highest Degree Offered

In terms of other clustered areas, the Systems/Networking and Software Engineering clusters are of highest interest to BS institutions and lowest interest to PhD100 institutions. Robotics/CPS is of most interest to PhD-granting institutions. Other CS topics are of most interest to BS institutions. Human-Computer Interaction is of most interest to the PhD-granting institutions. Computational Science is of most interest to the PhD100 institutions, while Compilers and Programming Languages is of most interest to the MS institutions.

Finally, Figure 7 shows the longitudinal percentage of positions devoted to areas that are more and some amount of interdisciplinary in nature. The 2018 results show that the impact of interdisciplinary areas is even more pronounced for PhD institutions with at least 32% and up to 65% of all positions devoted to these areas. These percentages are smaller than last year for these institutions. In contrast, MS and BS institutions have a range of 22-54% of positions are devoted to interdisciplinary areas. These percentages are increased from last year for MS and decreased for BS institutions.

5 Summary and Future Work

This work uses the same methodology as work from previous years 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 of 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 406 institutions seeking to fill hundreds of tenure-track faculty positions in Computer Science. There is a 17% one-year (and 52% two-year and 82% three-year) increase in

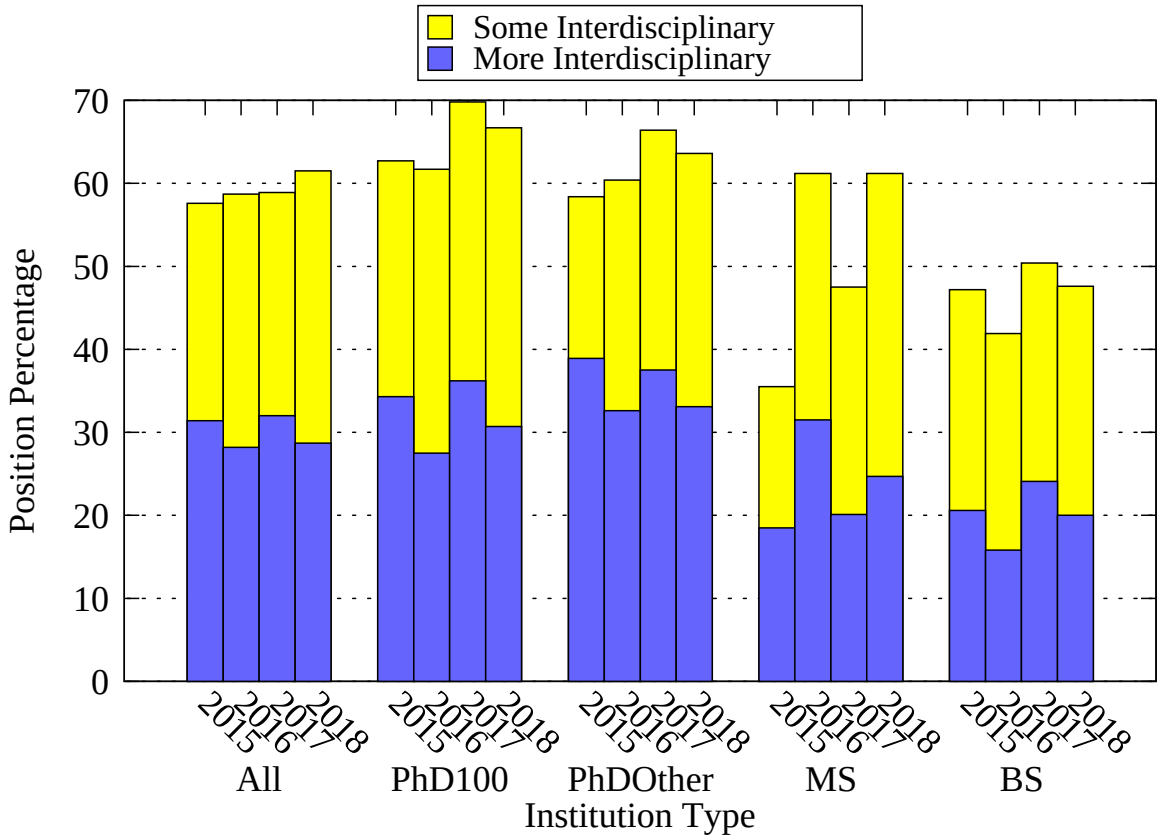

Figure 7: Investment in Interdisciplinary Clustered Areas by Institution Type Based on Highest Degree Offered

the number of institutions searching for tenure-track faculty in Computer Science and a 21% one-year (and 64% two-year and 107% three-year) increase in the number of positions being searched for. The number of institutions searching and positions seeking to be filled has increased the most for BS institutions.

In terms of specific areas, we found that the clustered areas of Security, Data Science and AI/Data Mining/Machine Learning are the areas of greatest investment. Aggregating the Data Science, AI/DM/ML and Databases clusters results in close to one-third of all hires sought in these data-oriented areas. We again 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, MS and BS institutions. Data Science is of most interest for other PhD institutions. 35% of positions for PhD institutions are in data-oriented areas. Finally, the abundance of potentially interdisciplinary areas is most pronounced for PhD institutions with 32-65% of all positions devoted to these areas.

A direction for future work is to examine how these searches translate into actual hires. Such follow-up was done for the first time in 2017 [4] and should continue to look both at numbers of actual hires as well as the areas in which these hires occur.

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`.

[2] Craig E. Wills. Analysis of current and future computer science needs via advertised faculty searches for 2016. *Computing Research News*, 28(1), January 2016. Full report at `http://web.cs.wpi.edu/~cew/papers/CSareas16.pdf`.

[3] Craig E. Wills. Analysis of current and future computer science needs via advertised faculty searches for 2017. *Computing Research News*, 29(1), January 2017. Full report at `http://web.cs.wpi.edu/~cew/papers/CSareas17.pdf`.

[4] Craig E. Wills. Outcomes of advertised computer science faculty searches for 2017. *Computing Research News*, 29(10), November 2017. Full report at `http://web.cs.wpi.edu/~cew/papers/outcomes17.pdf`.