Competitiveness Report: Computer Science Top Conference Performance Comparison and Collaboration in East Asia countries of China, Hong Kong, India, Japan, Korea, Singapore, Taiwan (2002-2006)

Hao-Hua Chu[1,2], Jane Yung-Jen Hsu[1,2], Polly Huang[2,3], Seng-Cho T. Chou[4], Ja-Ling Wu[1,2], Ming Ouhyoung[1,2], Ling-Jyh Chen[5], Tai-Yi Huang[6]

Department of Computer Science and Information Engineering[1]
Graduate Institute of Networking & Multimedia[2]
Department of Electrical Engineering[3]
Department of Information Management[4]
National Taiwan University, Taiwan
Institute of Information Science, Academia Sinica, Taiwan[5]
Department of Computer Science, National Tsing Hua University, Taiwan[6]

Abstract

In recent years, many Asian regions are busy implementing their large-scale academic research initiatives. This paper surveyed the Asian regions of China (Mainland), Hong Kong, India, Japan, Korea, Singapore, and Taiwan, and collected extensive statistics on their performance in some of most selective conferences in computer science. We then try to answer the following questions. How well are Asian researchers performing in comparison to the West? Where are Asian researchers strong/weak subjects in computer science? How much collaboration exists across different Asian regions and with the West?

Background

Recently, many Asian governments have launched large-scale academic programs, supported by public funding with scale in the U.S. billion dollars, to bring the quality of their top research universities to the international "first-rated" level (e.g., equivalent to MIT, Stanford, etc.). Given their importance, computer science and its related information technology (IT) are often placed as target areas for special development in these programs. For examples, Taiwan is currently implementing the "March towards First Rate Universities" program [4] with 5 years NTD 50 billion (about U.S. $1.6 billion) of special funding. This program's primary goal is to push its top research universities into the world's top 100 universities. In Mainland China, "Project 211" [5] is an ambitious project to construct some 100 universities. This is accompanied by another "985 project" that channels additional funding for building world-class universities in China. For example, two top universities in China, namely Peking and Tsing Hua Universities, both received 1.8 billion yuan (about U.S. $200 million) of special funding over a three years period. In Korea, "Brain Korea 21 (BK21)" [6] is a 7 years 1.4 trillion won (about U.S. $1.2 billion) program to develop world-class research universities with funding concentrating on certain academic fields and graduate schools. Its first phase ended in 2005 and has been followed by its second 7 years phase with even more ambitious funding scale of 2.8 trillion won (about U.S. $3 billion). In Japan, the "21 Century COE (Center of Excellence)" program [7] started in 2001, has provided yearly grants up to several U.S. million dollars for each of several hundred research centers that can demonstrate research excellence. In Singapore, universities have aggressively pursued international collaborations in research and teaching [8], establishing programs such as "Singapore-MIT alliance", the strategic partnership between Singapore Management University and CMU in information systems, and others.

This article hopes to offer a glimpse into how much success these academic programs have achieved in bringing quality of their research institutions to the international first-rated level in the field of computer science. This study was conducted at the end of 2006 to statistically *assess* and *compare* the research performance of different Asian regions in computer science. Our surveyed Asian regions include China (Mainland), Hong Kong, India, Japan, Korea, Singapore, and Taiwan. Since these academic programs normally come with funding sources for establishing international collaboration, this study also looks into how much *collaboration* are there among researchers across different Asian regions and with the West. Specifically, we try to answer the following three questions:

- How well do Asian researchers perform at some of the most selective research publications in different computer science subjects?
- How well do different Asian regions perform under different computer science subjects?
- How much collaboration exists across different Asian regions and with the West?

Method

First of all, we must first define what we mean by "first-rated level research publications". This requires defining a *benchmark* of publications for evaluating research performance. Since many researchers from acclaimed U.S. and European research institutions believe that in computer science, research conferences are the primary means of publishing and communicating research results [1][2]. Therefore, this study adapts their viewpoints. To come up with our benchmark publications, we have asked the opinions of experts in their respective research subjects and selected some 31 well-established research conferences in artificial intelligence, computer architecture, databases, media and applications, programming language, systems and networking, software engineering, and theory. Statistics are collected from papers published in these research conferences over the most recent 5 years (2002~2006). Our benchmark conferences are shown in Table 1. *Please note that the goal of defining our benchmark conferences is not to come up with a comprehensive listing of all top conferences covering all computer science subjects, but simply to provide a sufficient sample size for our analysis.*

For each benchmark conference listed in Table 1, we collect the following statistics for each year from 2002~2006.

- **Accept**: the total number of accepted papers.
- **Accept%**: the acceptance rate, which indicates the *selectivity* of a conference.
- **US25** *(or US25 papers)*: the number of accepted papers with at least one author(s) affiliated with the top 25 U.S. computer science departments using rankings from the U.S. News and World Report [3].
- **US25%**: the percentage of *US25* papers (defined previously) over all accepted papers. It indicates the *quality* of a research community participating in this conference.
- **Asia** *(or Asian papers)*: the number of accepted papers with at least one author(s) affiliated with a research institution *located* within our surveyed Asian regions. A paper with an author from a research institution located in Asia but a subsidiary of a U.S. or European company is counted as an Asian paper. However, a paper with an author from a research institution located in the U.S. or Europe but is a subsidiary of an Asian company is not counted as an Asian paper.
- **Asia%:** the percentage of *Asian* papers (defined previously).
- **TW, KR, JP, CN, HK, IN, SG:** the number of accepted papers with at least one author(s) affiliated with an institution *located* in our surveyed Asian regions of Taiwan (TW), Korea (KR), Mainland China (CN), Hong Kong (HK), IN (India), and SG (Singapore).
- **TW%, KR%, JP%, CN%, HK%, IN%, SG%:** the percentage of accepted papers from an Asian region over all accepted papers.
- **CoA** (*collaboration across Asian regions*): the number of accepted *Asian papers* that involve collaboration across two or more Asia regions. A qualified *CoA paper* must have its listed authors affiliated with at least two different Asia regions.
- **CoX** (*collaboration outside Asian regions*): the number of accepted *Asian papers* that involve collaboration with a research institution *outside* Asia. A qualified *CoX paper* must have some listed author(s) affiliated with an institution located in Asia and some author(s) affiliated with an institution in the West.
- **CoA%, CoX%**: the percentage of accepted CoA/CoX papers over all accepted Asian papers.

Results

We present our findings according to the three questions raised earlier.

How well do Asian researchers perform in our benchmark conferences?

Table 1 shows the collected statistics of our benchmark conferences from 2002 to 2006. The average acceptance ratio is 16%, or about one paper accepted out of six submitted papers. The average percentage of *US25* papers is 49%, or about half of accepted papers have listed author(s) affiliated with the top 25 U.S. universities in computer science. The low acceptance percentage and high *US25* percentage show some selectivity and quality of our benchmark conferences. Although Asian paper contribution constitutes only 8% of all accepted papers in these conferences, the yearly data in Table 2 shows a general upward trend from 7.2% (2002) to 9.8% (2006). Specifically, China, Singapore, and Hong Kong have exhibited faster moving trends in paper contributions over the past five years: China from 0.6% to 2.4%, Singapore from 0.9% to 1.8%, and Hong Kong from 1.1% to 2.8%. Figure 1 plots the average percentage of paper contribution from different Asian regions: China leads with 1.9%, followed by Japan (1.8%), Singapore (1.5%), Hong Kong (1.4%), India (0.7%), Korea (0.6%), and Taiwan (0.2%). Figure 2 plots the accumulative number of paper contribution from different Asian regions. China leads with 177 papers, followed by Hong Kong (154), Japan (137), Singapore (134), India (71), Korea (58), and Taiwan (35). Note that percentage of accepted paper contribution gives different ranking order than the total count of accepted paper contributions. In the former ranking metric, a paper at a conference (e.g., *SIGCOMM*) accepting 25 papers gives more weight than a paper at a conference (e.g., *INFOCOM*) accepted 250 papers.

Some readers may find this result surprising, given the impression that Japan is the technology leader in Asia, followed by the Asian four little dragons. Why is China on the top of the list? A part of the answer is Microsoft Research Asia (MSRA) in Beijing. According to MIT Technology Review, June 2004, MSRA is considered as world's hottest computer lab. Our collected statistics show a significant number of papers from China are authored or co-authored by MSRA. For examples, in computer graphics and information retrieval conferences of *SIGGRAPH/SIGIR* in 2005/2006, MSRA authored or co-authored 27 papers in 2005/2006, out of a total of 29 papers from China in these two years. Our statistics also show that MSRA collaboration extends beyond Mainland China into Hong Kong and Singapore, however, at a smaller scale.

How well do different Asian regions perform at different computer science subjects?

Table 3 shows statistics collected from our benchmark conferences categorized under eight computer science subjects: artificial intelligence, computer architecture, database, programming languages, software engineering, systems & networking, media & applications, and theory. Our statistics show that Asian researchers are better at the subjects of media & applications (Asian paper percentage 16.9%), database (11.8%), artificial intelligence (10.1%), and software engineering (8.1%); however, are weaker at the subjects of programming languages (5.5%), systems & networking (4.6%), theory (3.9%), and computer architecture (2.2%).

Mainland China excels in the subjects of media & applications (contributing 5.2% of accepted papers in 2002~2006) and artificial intelligence (3.3%), with impressive presence in *CVPR* (6%), *SIGGRAPH* (5.2%), *SIGIR* (9.2%), and *Multimedia* (12.7%). Japan demonstrates strength in the subjects of media & applications (4.4%) and programming languages (2.4%), with noticeably presence in *CHI* (4.7%), *UBICOMP* (6.8%), and *Multimedia* (7.1%). Singapore performs well across multiple subjects including database (3.2%), software engineering (3.1%), artificial intelligence (2.7%), and media & applications (2.6%), with noticeable presence in *SIGMOD* (6.3%), *SIGIR* (4.3%), and *Multimedia* (9.2%). Hong Kong is respectable in the subjects of database (4.0%) and artificial intelligence (2.5%), with significant presence at *SIGMOD* (6.4%) and *VLDB* (5.1%). India stands out in the subject of database (2.4%) with consistent presence at *VLDB* (3.1%) and *MOBICOM* (3.1%). Korea performs well at *SIGMOD* (3.3%).

How much collaboration exists across different Asian regions and with the West?

The *CoA/CoX* columns in Tables 2 and 3 show collaboration statistics in our benchmark conferences. About one third (32%) of the Asian papers involve collaboration with institutions from the West. A much smaller percentage (9%) involves collaboration among institutions across Asian regions. In other words, when Asian researchers look for research partners, collaboration statistics show a *preference* to collaborate with the West rather than their Asian neighbors. Why is there a collaboration preference? There could be several explanations. One explanation is that these conferences are in English, having

an English co-author improves paper writing and presentation quality. The second explanation could be that a significant portion of Asian researchers had worked or been educated in the West, and they are continuing that working relationship. The third explanation could be that working with researchers in the elite Western institutions (e.g., MIT, Stanford, etc.) results in a higher chance of acceptance in these selective conferences.

Table 2 shows increasing trends for both Asian regional collaboration (*CoA*) papers and Western collaboration (*CoX*) papers: *CoA* from 4 (2002) to 24 (2006), and *CoX* from 31 (2002) to 55 (2006). Table 3 shows the collaboration statistics under different computer science subjects. Asian researchers are most active in collaboration among themselves at the subjects of theory (consisting of 13% of accepted Asian papers), artificial intelligence (11%), media & applications (10%), systems & networking (10%), and database (9%). Since three of these subjects are also Asian stronger subjects, this indicates some presence of an Asian research community in these three subjects. On the other hands, Asian researchers in theory (53%), programming languages (50%), database (44%), and systems & networking (42%) are most active in collaboration with the West. Since three of these subjects are Asian weaker subjects, collaboration with the West seems to make sense as it provides good learning opportunities for Asian researchers.

Discussion

It is encouraging that many governments in Asia have realized the importance of academic research and have recently committed a larger share of public funding in academic research programs. Our statistics show some effectiveness in these programs in Asia, particularly in China, Singapore, and Hong Kong. We also expect to continue to see an upward trend in the percentage of Asian papers at these benchmark conferences in the near future. In many subjects of computer science (artificial intelligence, database, and media & applications), the past five years has been a noticeable period when Asian researchers are catching up in research quality with the West. We are also hopeful of an improved Asian paper presence at subjects traditionally weak in Asia in the past, such as computer architecture, systems & networking, and theory.

We comment that our results are not normalized with the number of universities and researchers in each Asian region. If our results are normalized with the number of research universities in each Asian region, Singapore (about 4 research universities) and Hong Kong (about 8) with smaller population sizes would stand out even further. At the same time, Korea (number in tens) and Taiwan (number in tens) would come closer to Japan (number in hundreds). China (number in hundreds) would also seem less dominant.

For Taiwan's relatively weak performance in the benchmark conferences, we offer the following explanation. There exists an academic legacy that places higher emphasis on the journal publications based on the Science Citation Index (SCI®) [9] over conference publications. This preference has somehow been supported by the much publicized "world university rankings" report [10] from Shanghai Jiao Tong University, which counts SCI journal publications. However, we are hopeful that adjustments in Taiwan's evaluation system will be introduced to better conform to the current international academic trend in computer science research.

Collaboration statistics show that Asian regional collaboration produces much less papers than collaboration with the West. Although collaboration with reputable Western researchers is a short path to produce quality research work that can meet the standard of these benchmark conferences, however, considering the long run is beneficial to breed an active research community locally in Asia that can work together to produce equal quality papers to those with collaboration from the West. Although our collaboration statistics shows some presence of this Asian research community at some subjects which Asian researcher are good at, this collaboration effort can certainly enlist more active participation and encouragement in Asia. We are hopeful that through both *competition* and *collaboration*, Asian regions can effectively implement these large-scale academic research programs and succeed in achieving the "international first-rated" research quality.

References

[1] Patterson, D., The health of research conferences and the dearth of big idea papers. Communi-

cations of the ACM, Vol. 47, No. 12, December 2004.
[2] Patterson, D., Snyder, L., and Ullman, J. Evaluating computer scientists and engineers for promotion and tenure. Computing Research News, September 1999.
[3] "Amercica' Best Graduate Schools 2006, Computer Science (Ph.D.) ", U.S. News & World Report website.
[4] National Taiwan University, "March Towards Top University", http://top100.ntu.edu.tw/.
[5] China Education and Research Network, "Project 211: A Brief Introduction", http://www.edu.cn/HomePage/english/education/highedu/211/index.shtml.
[6] Korean Ministry of Education and Human Resources Development, The second stage of BK21, fostering a world class talent, http://www.bk21.or.kr/.
[7] Japan Society for the Promotion of Science, 21st century COE program – targeted support for creating world-standard research and education bases (centers of excellence), http://www.jsps.go.jp/english/e-21coe/index.html.
[8] National Singapore University, International collaboration in research and teaching, http://www.nus.edu.sg/corporate/global/global_partners.htm.
[9] Thompson Corporation, "Science Citation Index ®" http://scientific.thomson.com/products/sci/.
[10] Shanghai Jiao Tong University, "Academic Rankings of Worlds' Universities", http://ed.sjtu.edu.cn/rank/2006/ARWU2006FULLLIST-BY%20RANK%20(PDF).pdf

Table 1. Statistics of our benchmark conferences (2002~2006). It shows average number, except the **Accept** column counts the total number of papers. Note only regular/oral papers (i.e., excluding poster, short, special track, and invited papers) are counted.

Subjects	Conferences	Accept	Accept%	US25%	Asia%	TW%	KR%	JP%	CN%	HK%	IN%	SG%
Artificial intelligence (4)	*AAAI*	440	18%	43%	8.0%	0.0%	0.0%	1.9%	1.1%	2.8%	0.0%	2.9%
	IEEE ICCV	88	4%	40%	8.0%	0.0%	0.0%	3.5%	3.4%	0.0%	0.0%	1.2%
	IEEE CVPR	244	6%	39%	12.1%	0.0%	0.4%	1.6%	**6.0%**	3.6%	0.0%	2.0%
	ACM SIGKDD	220	12%	35%	10.7%	0.0%	0.0%	0.9%	3.7%	2.2%	1.8%	3.0%
Computer architecture (4)	*ACM ICSA*	170	17%	74%	1.2%	0.0%	0.0%	1.1%	0.0%	0.0%	0.0%	0.0%
	IEEE HPCA	137	18%	53%	3.7%	0.0%	0.7%	2.1%	0.0%	0.0%	0.8%	0.0%
	IEEE/ACM MICRO	171	22%	62%	1.6%	0.0%	0.1%	0.7%	0.0%	0.0%	1.0%	0.0%
	ACM ASPLOS	86	17%	79%	2.3%	0.0%	0.0%	2.3%	0.0%	0.0%	0.0%	0.0%
Database (3)	*ACM SIGMOD*	287	16%	50%	18.2%	0.0%	3.3%	1.1%	0.7%	**6.4%**	2.2%	**6.3%**
	ACM PODS	148	19%	34%	3.7%	0.0%	0.0%	0.6%	0.0%	0.6%	1.9%	0.6%
	VLDB	434	17%	35%	13.6%	0.5%	2.1%	0.4%	0.9%	**5.1%**	3.1%	2.7%
Programming languages (2)	*ACM POPL*	145	19%	39%	3.3%	0.0%	1.3%	1.4%	0.6%	0.0%	0.0%	0.0%
	ACM PLDI	145	20%	54%	7.7%	0.0%	0.0%	3.4%	2.8%	0.0%	0.7%	0.8%
Software engineering (2)	*IEEE ICSE*	228	13%	26%	7.7%	0.0%	0.5%	1.9%	0.3%	1.5%	0.0%	3.5%
	ACM FSE	132	17%	35%	8.5%	0.6%	0.0%	0.0%	2.0%	3.0%	0.8%	2.7%
Systems & networking (8)	*ACM SIGCOMM*	154	10%	71%	1.7%	0.5%	0.0%	0.5%	0.6%	0.0%	0.0%	0.0%
	IEEE INFOCOM	1143	19%	44%	9.2%	1.6%	1.2%	0.7%	1.8%	2.9%	1.1%	1.0%
	ACM MOBICOM	137	9%	52%	10.8%	0.0%	1.9%	1.5%	1.5%	1.3%	3.1%	2.3%
	ACM SOSP	42	15%	88%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	USENIX/ACM OSDI	81	16%	70%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	ACM SIGMETRICS	143	13%	54%	4.8%	0.0%	1.3%	0.7%	0.0%	2.1%	0.6%	0.0%
	IEEE SECURITY & PRIVACY	108	14%	60%	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.6%
	ACM CCS	173	15%	49%	4.5%	0.6%	0.0%	0.6%	0.0%	0.0%	0.0%	1.8%
Media & applications (6)	*ACM SIGGRAPH*	415	19%	56%	15.6%	0.4%	1.0%	2.3%	**5.2%**	2.5%	0.0%	0.2%
	ACM SIGCHI	440	19%	43%	14.1%	0.0%	0.4%	**4.7%**	0.9%	0.2%	0.0%	0.0%
	ACM SIGIR	293	19%	33%	18.3%	0.3%	2.4%	1.9%	**9.2%**	3.2%	0.0%	**4.3%**
	ACM Multimedia	241	16%	24%	31.3%	2.1%	1.1%	**7.1%**	**12.7%**	0.9%	0.0%	**9.2%**
	ACM UBICOMP	134	14%	36%	7.5%	0.0%	0.7%	**6.8%**	0.0%	0.0%	0.0%	0.0%
	W3C WWW	375	14%	26%	14.3%	0.8%	0.0%	3.7%	3.5%	2.1%	2.9%	2.1%
Theory(2)	*ACM FOCS*	404	29%	53%	4.1%	0.3%	0.0%	1.0%	0.2%	1.7%	1.2%	0.0%
	IEEE STOC	335	25%	55%	3.7%	0.0%	0.0%	0.3%	1.4%	0.9%	2.0%	0.3%
Average (Total)		**(7693)**	**16%**	**49%**	**8.0%**	**0.2%**	**0.6%**	**1.8%**	**1.9%**	**1.4%**	**0.7%**	**1.5%**

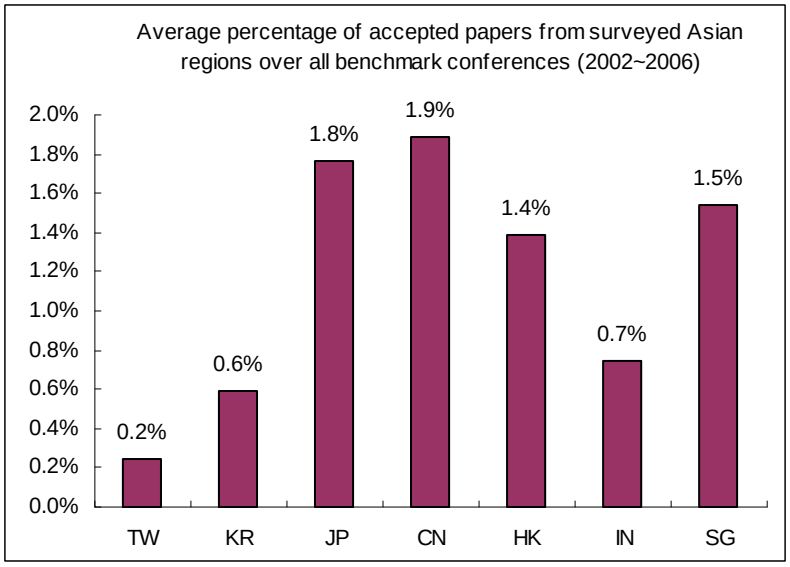

Figure 1. Percentage of paper contribution from different Asian regions in our benchmark conferences (2002~2006).

Table 2. Asian papers in the benchmark conferences (2002~2006)

Years	Asia(%)	TW(%)	KR(%)	JP(%)	CN(%)	JP(%)	HK(%)	IN(%)	SG(%)	CoA(%)	CoX(%)
2002	103(*7.2%*)	5(*0.1%*)	9(*0.5%*)	29(*2.1%*)	16(*0.6%*)	29(*2.1%*)	18(*1.1%*)	18(*1.2%*)	9(*0.9%*)	4(*4%*)	31(*30%*)
2003	123(*6.9%*)	10(*0.5%*)	8(*0.7%*)	26(*1.7%*)	26(*1.3%*)	26(*1.7%*)	20(*0.9%*)	12(*0.5%*)	26(*1.5%*)	5(*4%*)	32(*26%*)
2004	144(*6.9%*)	8(*0.1%*)	14(*0.5%*)	24(*1.0%*)	37(*1.9%*)	24(*1.0%*)	39(*1.5%*)	9(*0.6%*)	30(*1.6%*)	17(*12%*)	49(*34%*)
2005	162(*8.7%*)	6(*0.1%*)	7(*0.3%*)	32(*1.7%*)	47(*2.2%*)	32(*1.7%*)	33(*1.4%*)	15(*1.0%*)	38(*2.3%*)	17(*11%*)	60(*37%*)
2006	172(*9.8%*)	6(*0.1%*)	20(*0.8%*)	26(*1.5%*)	50(*2.4%*)	26(*1.5%*)	44(*2.8%*)	17(*1.1%*)	31(*1.8%*)	24(*14%*)	55(*32%*)
Total	**704**	**35**	**58**	**137**	**176**	**137**	**154**	**71**	**134**	**67(*9%*)**	**227(*32%*)**

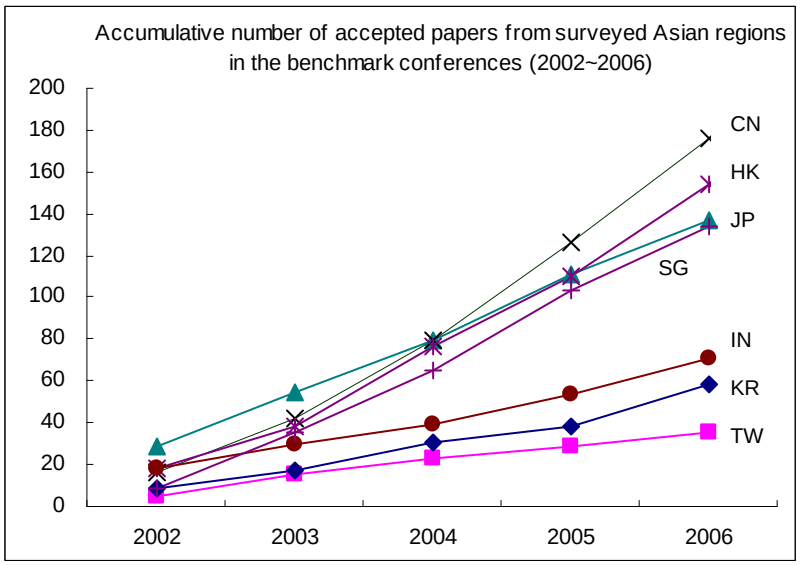

Figure 2. Accumulative number of papers from our surveyed Asian regions in the benchmark conferences (2002~2006).

Table 3. Benchmark conference statistics (2002~2006) categorized under eight computer science subjects. Each entry shows total paper count (the average paper contribution) from each Asia region. Highlighted numbers (> 2.5%) show good performance.

Subjects	Accept%	US25%	Asia(%)	TW(%)	KR(%)	JP(%)	CN(%)	HK(%)	IN(%)	SG(%)	CoA(%)	CoX(%)
Artificial intelligence	11%	39%	97(*10.1%*)	0(*0.0%*)	1(*0.1%*)	20(*1.6%*)	**32(*3.3%*)**	**25(*2.5%*)**	4(*0.7%*)	**25(*2.7%*)**	11(*11%*)	23(*24%*)
Computer architecture	19%	66%	12(*2.2%*)	0(*0.0%*)	1(*0.3%*)	8(*1.3%*)	0(*0.0%*)	0(*0.0%*)	3(*0.5%*)	0(*0.0%*)	0(*0%*)	2(*17%*)
Database	17%	40%	116(*11.8%*)	2(*0.2%*)	17(*1.8%*)	6(*0.7%*)	6(*0.5%*)	**41(*4.0%*)**	22(*2.4%*)	**32(*3.2%*)**	10(*9%*)	51(*44%*)
Programming languages	20%	49%	16(*5.5%*)	0(*0.0%*)	2(*0.6%*)	**7(*2.4%*)**	5(*1.7%*)	0(*0.0%*)	1(*0.4%*)	1(*0.4%*)	0(*0%*)	8(*50%*)
Software engineering	15%	31%	29(*8.1%*)	1(*0.3%*)	1(*0.2%*)	4(*1.0%*)	4(*1.2%*)	7(*2.2%*)	1(*0.4%*)	**12(*3.1%*)**	1(*3%*)	6(*21%*)
Systems & networking	14%	58%	140(*4.6%*)	20(*0.4%*)	19(*0.6%*)	13(*0.6%*)	24(*0.6%*)	39(*0.9%*)	17(*0.7%*)	19(*0.8%*)	14(*10%*)	59(*42%*)
Media & applications	17%	36%	261(*16.9%*)	11(*0.6%*)	17(*0.9%*)	**74(*4.4%*)**	**99(*5.2%*)**	32(*1.5%*)	11(*0.5%*)	**44(*2.6%*)**	27(*10%*)	62(*24%*)
Theory	27%	54%	30(*3.9%*)	1(*0.1%*)	0(*0.0%*)	5(*0.7%*)	6(*0.8%*)	10(*1.3%*)	12(*1.6%*)	1(*0.2%*)	4(*13%*)	16(*53%*)