

Mulheres na Computação: Análises por Sub-Áreas

Bárbara D. Duarte, Ana Flávia C. Moura, Mirella M. Moro

Universidade Federal de Minas Gerais

{barbaradm, anaciriaco, mirella}@dcc.ufmg.br

***Resumo.** Este artigo apresenta a presença feminina nas sub-áreas da Computação brasileira a partir da participação em comitês de programa de simpósios da Sociedade Brasileira de Computação ao longo do tempo.*

***Abstract.** This paper characterizes the female participation in CS subareas through the program committees of the main SBC symposia over time.*

1. Introdução

Este artigo propõe uma metodologia diferenciada para caracterizar diversidade de gênero em sub-áreas da Computação: a presença feminina nos comitês de programa (CP) de simpósios (SBs) promovidos pela SBC. A justificativa é que os CPs são uma amostra da parte mais ativa e consolidada da comunidade de uma sub-área. Autores dos artigos dos SBs também formam uma parte relevante de tais comunidades. Porém, já existem estudos que caracterizam tais autores, e.g. [1] e [3]; e, muitas vezes, não existe diferenciação entre autores “de um artigo só” (estudantes) e autores presentes nas comunidades ao passar do tempo. Além disso, membros de um CP têm poder de decisão sobre quais artigos são publicados, então possuindo presença mais decisória na comunidade.

A partir de dados provenientes dos comitês de programa de 17 eventos nacionais, as questões que conduzem este estudo são: como é a presença feminina por evento e como a mesma evolui no tempo; e como as mulheres participam de comitês de eventos diferentes. Iniciativas brasileiras para atração de meninas para a Computação podem se beneficiar dos resultados aqui apresentados explorando melhor as áreas que possuem mais profissionais mulheres. Este artigo segue com trabalhos relacionados e nossas contribuições, a metodologia, as avaliações e discussões sobre resultados.

2. Trabalhos Relacionados

Existem muitas iniciativas para atrair mulheres para a Computação. Várias iniciativas nacionais são apresentadas anualmente no WIT durante o Congresso da SBC. Entre as internacionais, citamos o *SAILORS - Stanford Artificial Intelligence Laboratory's Outreach Summer*, como uma ação diferenciada por focar em projetos práticos de uma **sub-área** específica da Computação, a Inteligência Artificial [4]. A questão que surge então é: quais seriam as **outras** sub-áreas da Computação mais convidativas para atrair meninas? Alguns estudos realizados com meninas em idade escolar quantificam o interesse acerca das sub-áreas, como [2]. No entanto, o grau de conhecimento é uma variável determinante dos seus resultados. Isto é, se o nível de compreensão acerca de uma determinada sub-área fosse maior, essa não se tornaria mais atrativa? Neste artigo, caracterizamos as diversas áreas no contexto nacional para possibilitar a identificação de aspectos que possam contribuir para essa atratividade. A metodologia é diferenciada por considerar os comitês de programas dos eventos científicos, ao invés de os autores que publicam nos mesmos e que são avaliados em vários estudos existentes, como por exemplo [1] e [3].

Tabela 1. Áreas e eventos nacionais considerados com palavras-chave

Evento	Simpósio Brasileiro em: Algumas palavras-chave relevantes (além do nome do simpósio)
BRACIS	Intelligent Systems: inteligência artificial e computacional, aprendizado de máquina e redes neurais
ETC	Teoria da Computação: fundamentos da computação, algoritmos, combinatória e otimização
IHC	Fatores Humanos em Sist. Computacs.: soluções computacionais práticas, ética e moral no desenvolvimento de tecnologias
SBBB	Bancos de Dados: projeto e otimização de banco de dados, métodos de acesso, consultas e indexação
SBCM	Computação Musical: musicologia computacional, processamento de som e desempenho multimídia
SBES	Software: Teoria e Prática: desenvolvimento, teste e reutilização de software, linguagens de programação
SBESC	Eng. de Sistemas Computacionais.: sistemas operacionais, de tempo real, embarcados e críticos
SBIE	Informática na Educação: informática educacional, tecnologia no ensino e ensino à distância
SBMF	Métodos Formais: métodos formais, verificação de sistemas e linguagens formais
SBRC	Redes Computs. e Sist. Distribuídos: protocolos de computadores, computação móvel e sistemas distribuídos
SBSC	Sistemas Colaborativos: sistemas de compartilhamento, cooperativos e de comunicação
SBSeg	Segurança Info. e Sist. Computacs.: segurança da informação, segurança de sistemas, privacidade e criptografia
SBSI	Sistemas de Informação: fundamentos de sistemas de informação, processos de negócios e planejamento estratégico
SIBGRAP	Graphics, Patterns and Images: computação gráfica, visão computacional, processamento de imagens e padrões
SVR	Virtual and Augmented Reality: realidade virtual, realidade aumentada, simulação virtual e ambiente virtual
WebMedia	Sistemas Multimídia e Web: mídia, hipermídia, web, TV digital e redes sociais
WSCAD	Sist. Computacs. de Alto Desempenho: arquitetura de computadores, processamento de alto desempenho e sistemas distribuídos

3. Metodologia

O objetivo desta pesquisa é apurar a participação de mulheres nas diferentes áreas de Computação. Para tal, utilizou-se a seguinte metodologia.

Definição de Eventos. Considera-se um conjunto dos eventos nacionais das sub-áreas de Computação realizados pela Sociedade Brasileira de Computação¹, conforme Tabela 1. Os seguintes eventos não são incluídos por não possuírem informações suficientes online: S.B. Qualidade de Software, S.B. Games e Entretenimento Digital, S.B. Concepção de Circuitos e Sistemas Integrados, e S.B. Robótica. Também não foi encontrada informação sobre BRACIS e WSCAD em 2011; e o SBSC foi finalizado em 2017.

Coleta de Dados. A coleta dos dados foi realizada de maneira manual, pois eventos não possuem uma forma unificada de armazenamento. A informação mínima é: nome/sigla do evento, ano, e nome dos participantes do CP, no período estudado de 2011 a 2018.²

Gênero. Nenhum evento oferece informação sobre gênero dos integrantes do seu CP. Então, realizou-se uma classificação automática com uma biblioteca de processamento de linguagem natural.³ Para os casos não identificados, foi utilizada outra biblioteca específica para identificação de gênero.⁴ Após utilizar tais bibliotecas, cerca de 30% das pessoas ainda estavam com gênero sem identificação (por exemplo, nomes como Ariel que servem para homens e mulheres). Assim, tais nomes foram manualmente classificados com a busca das respectivas pessoas online (considerando universidade e país).

Desambiguação. Esta etapa agrupa registros diferentes para a mesma pessoa de acordo com primeiro e último nome (e.g, Mirella M. Moro e Mirella Moro são a mesma pessoa).

Estatísticas. A base inicial contém 10.489 registros. Após os processos supracitados, restaram 3.382 registros, que representam o número de pessoas nos CPs, divididos nos eventos conforme a Tabela 2. Dois eventos foram excluídos após a coleta: o ETC com apenas três edições e o bi-anual SBCM, que podem ser melhor analisados isoladamente.

¹SBC Simpósios Realizados: <http://www.sbc.org.br/eventos/eventos-realizados>

²Essas são as edições para as quais conseguimos todos os dados disponíveis online; em um passo seguinte, planeja-se coletar os dados nos anais publicados fisicamente ou disponibilizados em mídia digital, o que necessita de um trabalho manual ainda maior.

³Natural Language Toolkit: <http://www.nltk.org>

⁴Gender Guesser: <http://pypi.org/project/gender-guesser>

Tabela 2. Tamanho do comitê de programa por evento por ano

evento	2011	2012	2013	2014	2015	2016	2017	2018	min	máx	média
BRACIS	X	135	127	185	111	131	173	192	111	192	150.6
IHC	55	49	80	85	94	106	138	120	49	138	90.9
SBBD	52	43	59	44	48	50	52	51	43	59	49.9
SBES	84	85	76	78	33	31	84	45	31	85	64.5
SBESC	70	60	85	46	34	38	70	45	34	85	56.0
SBIE	153	217	195	222	149	226	124	244	149	244	191.2
SBMF	35	40	42	39	41	47	46	56	35	56	43.3
SBRC	102	103	120	112	111	100	103	118	100	120	108.6
SBSC	41	51	48	44	52	48	42	X	41	52	46.6
SBSeg	91	42	47	49	46	56	56	65	42	91	56.5
SBSI	69	98	108	83	105	94	128	132	69	132	102.1
SIBGRAPI	155	59	60	268	105	146	153	156	59	268	137.8
SVR	81	89	82	83	80	35	56	57	35	89	70.4
WebMedia	48	67	62	72	77	96	96	98	48	98	77.0
WSCAD	X	86	85	80	90	77	41	78	41	90	76.7

Perguntas. A avaliação da composição por gênero dos comitês de programa dos eventos é guiada pelas seguintes perguntas: (1) Como a presença feminina varia de acordo com o evento? (2) Como a participação feminina evolui no somatório de todos os eventos? (3) Como é essa evolução nos eventos de maior e menor participação feminina?

4. Avaliações e Resultados

Esta seção apresenta e discute os resultados conforme as três perguntas definidas.

Presença Feminina por Evento. A maneira mais simples de averiguar a presença feminina é contar o número de mulheres em cada edição. Porém, os eventos possuem tamanhos variados; então tal número precisa ser acompanhado da porcentagem de mulheres para comparação mais uniforme. Outra questão importante é que alguns membros do CP participam de várias edições ao longo dos anos. Então, são dois números diferentes: o de mulheres em termos de **peçoas** que estão no comitê e o de **participações** de mulheres ao longo do tempo. Por exemplo, considerando o comitê *X*, uma mulher *A* participou nos anos $\{y_1, y_2, y_3\}$; o homem *B* participou no ano $\{y_1\}$ e o homem *C* participou nos anos $\{y_2, y_3\}$. O número de **peçoas** que participam desse comitê é três, com percentual de 33% mulheres; já o número de **participações** é seis, com percentual de 50% mulheres.

A Figura 1 apresenta o número total e percentual de **peçoas**, e a Figura 2 o somatório de **participações** nos comitês de programa separados por gênero no período. Com base nesses totais, os eventos SBIE, IHC e SBSI têm o maior número absoluto de mulheres em seus comitês, bem como participações de mulheres. Nos valores percentuais, o evento SBSC assume a terceira posição, colocando o SBSI em quarta. É importante notar que SBSC é um evento bem menor que os três supracitados. Sobre os eventos “mais masculinos”, em termos absolutos, os cinco eventos com menos pessoas e participações femininas são: SBESC, SVR, WSCAD, SBSeg e SBMF (com pequenas variações).

Os quatro eventos mais femininos possuem grande viés interdisciplinar. Especialmente a Informática na Educação, é notório que uma das áreas de atuação ditas tipicamente femininas no Brasil seja a Educação. Uma possível análise seria identificar a área de atuação primária ou de formação dos membros dos comitês de programa.

Evolução do Total de Participação Feminina. A Figura 3 ilustra o número de mulheres ao longo do tempo para o conjunto de eventos (Tabela 2). O gráfico da esquerda ilustra o número absoluto por gênero e o da direita a mesma informação mas em porcentagem, onde observa-se participação de 70% em média para homens e 30% de mulheres.

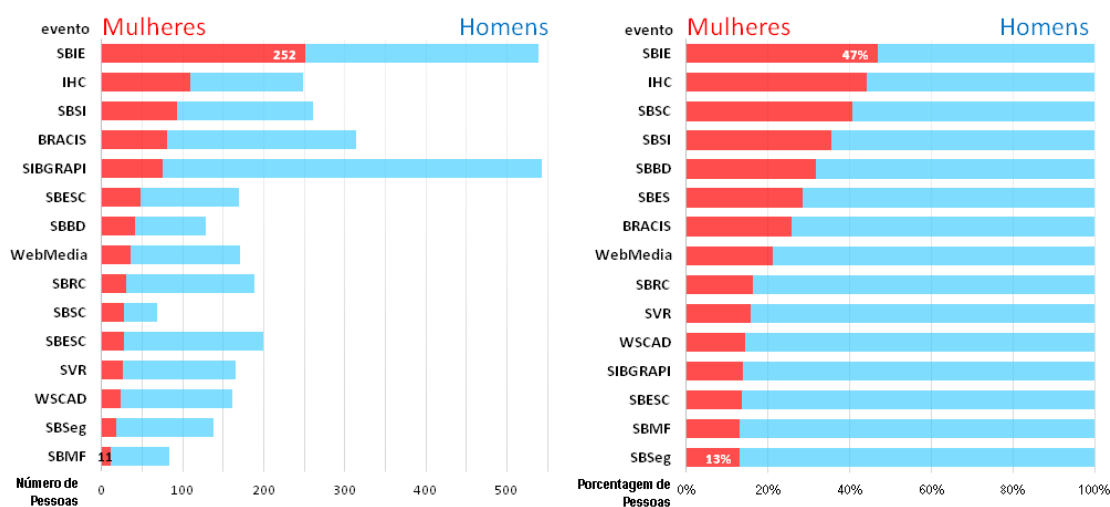


Figura 1. Números absoluto e percentual de pessoas por gênero (2011-2018).

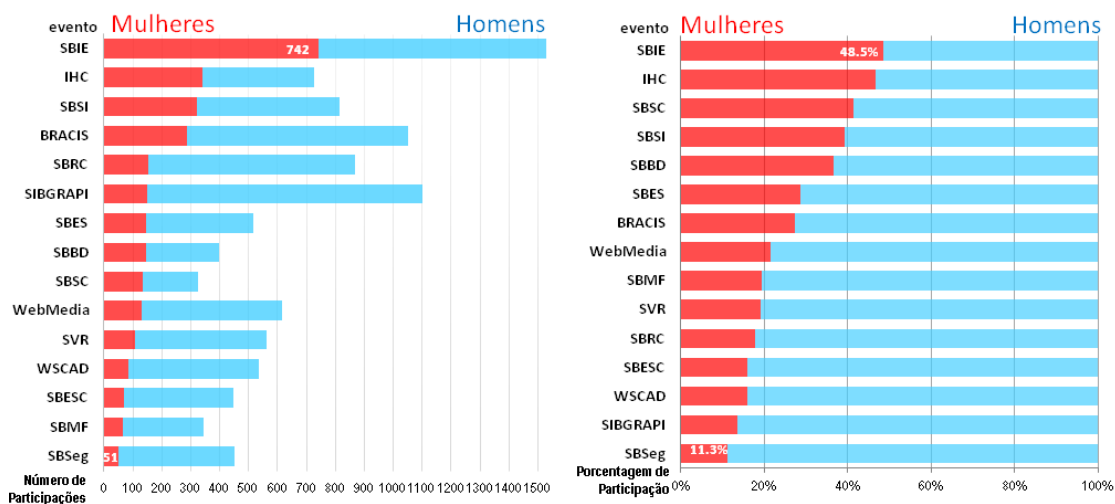


Figura 2. Números absoluto e percentual de participantes por gênero (2011-2018).

A presença de 30% de mulheres não acompanha o número de mulheres matriculadas na Computação, que caiu de 16.3% para 14% no mesmo período. Pode-se especular que a falta de modelos na área de Computação não seja problema, mas talvez a falta de divulgação sobre a presença feminina de 30% na área? Ou talvez essa taxa de 30% mulheres deva-se à interdisciplinaridade dos eventos mais femininos. Pode-se também utilizar a interdisciplinaridade a nosso favor, disseminando as oportunidade de trabalhar em diversas áreas de conhecimento aliados à Computação.

Evolução de Participação Feminina dos Eventos Extremos. A discussão anterior mostra que a participação feminina no Simpósio Brasileiro de Informática na Educação (SBIE) é a maior, e no Simpósio Brasileiro de Segurança da Informação e Sistemas Computacionais (SBSeg) é a menor. A Figura 4 detalha a evolução dessas participações em números absolutos à esquerda e percentagem à direita. A diferença entre esses extremos é clara, principalmente na percentagem. Percebe-se que a derivada atual para o SBSeg é positiva indicando um aumento da participação feminina. É possível observar o crescimento de participação feminina para outros eventos de áreas “mais masculinas” como SBESC, SVR e SBSeg, além de um movimento de queda em eventos de áreas “mais fe-

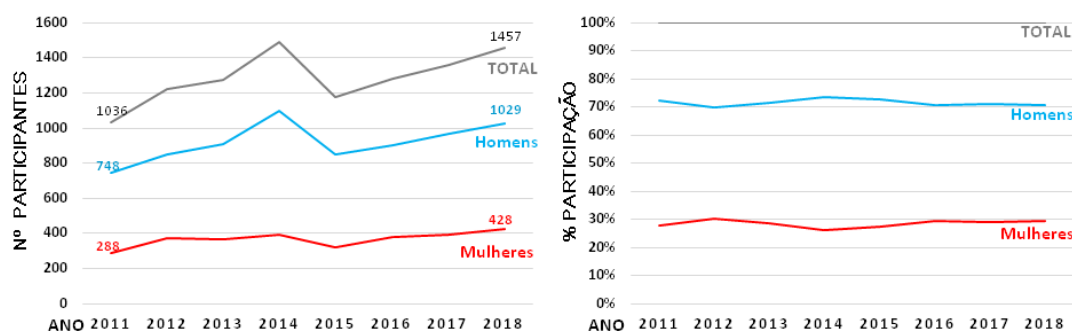


Figura 3. Evolução dos números absolutos e percentuais do total por gênero.

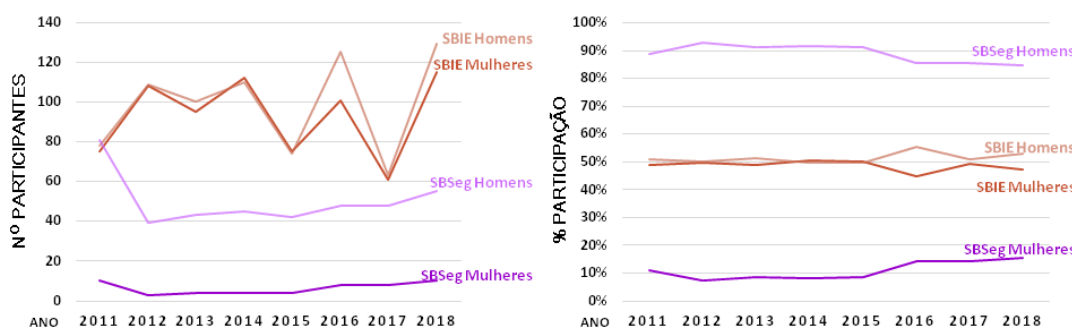


Figura 4. Números absoluto e percentual do total por gênero.

mininas” como SBIE, SBSI e SBSC. Destaca-se também o grande aumento de mulheres no IHC, e que talvez possa o tornar “mais feminino” do grupo analisado em médio prazo.

5. Considerações Finais

Este artigo apresentou a primeira etapa de um projeto sobre atração de mulheres para a Computação: caracterizar as sub-áreas através da análise da presença feminina nos comitês de programa dos simpósios promovidos pela SBC. O principal resultado é a presença maior de mulheres em áreas mais interdisciplinares, relacionadas à aplicação da computação ou com aspecto social claro. Como próximas atividades, pretende-se compreender melhor e explorar mais características das áreas femininas nos materiais e atividades a serem realizados em nossos projetos de atração de talentos para a área.

Agradecimentos. Trabalho parcialmente financiado por CNPq.

Referências

- [1] D. Arruda et al. Brazilian computer science research: Gender and regional distributions. *Scientometrics*, 79(3):651–665, 2009.
- [2] K. da Silva Figueiredo, R. Vitorassi, E. Monteiro, and S. de Oliveira Carneiro. Percepções de alunas de ensino médio sobre as subáreas da computação. In *WIT*. SBC, 2017.
- [3] L. H. C. de Lima, G. Penha, L. M. de Alencar Rocha, M. M. Moro, A. P. C. da Silva, A. H. F. Laender, and J. P. M. de Oliveira. The collaboration network of the Brazilian Symposium on Databases – 30 editions of history. *JBCS*, 23(10), 2017.
- [4] M. E. Vachovsky et al. Toward more gender diversity in CS through an artificial intelligence summer program for high school girls. In *SIGCSE*, pages 303–308, 2016.