

# Um Estudo Empírico sobre Crowdsourcing: O Brasil na plataforma TopCoder

Carlos H. Lopes<sup>1</sup>, Ricardo Rodrigo M. Melo<sup>1</sup>, Renato P. Ferreira<sup>1</sup>, Muller G. Miranda<sup>1</sup>, Cleidson De Souza<sup>1,2</sup>

<sup>1</sup>Programa de Pós Graduação em Ciência da Computação – Universidade Federal do Pará (UFPA) – Belém – PA – Brasil

<sup>2</sup>Instituto Tecnológico Vale – Belém – PA – Brasil

{carloshugolopes, ricardorodrigomm, renpina, mulgsm}@gmail.com,  
cleidson.desouza@acm.org

**Abstract.** *This paper describes an analysis of Brazilian software developers in the largest software engineering crowdsourcing platform, TopCoder. We compare Brazilian active coders in this platform with active coders from other countries, especially BRICS, Latin America and Western Europe. Our results can be used to further Brazilian participation in the platform.*

**Resumo.** *Este artigo descreve uma análise quantitativa dos programadores do Brasil que atuam na maior plataforma crowdsourcing do mundo na área de engenharia de software, a TopCoder. Além disso, considerando apenas os desenvolvedores ativos, uma análise é feita comparando os brasileiros com os programadores de outros países, assim como os brasileiros comparados com os programadores das regiões do BRICS, América Latina e Leste Europeu. Os resultados deste estudo podem ser utilizados para melhorar a participação dos engenheiros de software brasileiros.*

## 1. Introdução

Segundo Brabham (2008), o termo *crowdsourcing* foi cunhado por Howe e Robinson em Junho de 2006. Esse termo descreve um novo modelo de negócio baseado na *web* que aproveita os conhecimentos e soluções de pessoas geograficamente distribuídas, mas conectadas via Internet. Por exemplo, uma pessoa ou companhia divulga, por meio da Internet, um problema a ser resolvido. Um número vasto de pessoas oferecem soluções para o problema e a recompensa é dada para a solução vencedora. Neste contexto, ambos se beneficiam: tanto a empresa por obter a solução de seu problema ou produto rapidamente e com melhores custos financeiros, quanto o trabalhador que desenvolveu a solução ao ter a possibilidade de ganhar uma recompensa, seja em dinheiro, prêmios, ou até mesmo a própria fama (Malone et al. 2010). A InnoCentive é um exemplo de empresa que oferece prêmios para pesquisadores que apresentam soluções para um problema científico (Zatti 2013) (Malone et al. 2011).

O surgimento do *crowdsourcing* foi proporcionado pelo conceito de inteligência coletiva. Segundo (Alves et al. 2011), a inteligência coletiva é definida como uma inteligência presente em todas as partes do mundo, coordenada em tempo real, e que resulta em competências eficientes continuamente valorizadas. A multidão colabora em

um objetivo comum e pode até mesmo gerar resultados muito melhores do que os feitos por uma só pessoa. Esses conceitos foram potencializados com o advento das novas tecnologias de comunicação, como a Internet. Alguns pesquisadores (e.g., (Malone et al. 2011)) vêem o modelo do *crowdsourcing* como uma nova forma de como o trabalho é ou será feito pelas empresas. Em particular, Malone e colegas afirmam que uma nova era, a era da hiperespecialização, está se iniciando. Ele afirma ainda que muito das riquezas que o mundo está aproveitando atualmente vem do ganho da produtividade dividindo o trabalho em tarefas cada vez menores realizadas por trabalhadores cada vez mais especialistas. Hoje, isso ocorre graças ao aumento do conhecimento do trabalho e das tecnologias de comunicações.

Um exemplo baseado no modelo de *crowdsourcing* é a plataforma TopCoder que é disponibilizada pela empresa Appirio onde os clientes fornecem seus projetos para a TopCoder que separa os mesmos em módulos. Cada módulo é aberto para mais de 600.00 programadores inscritos na plataforma. O objetivo é que eles completem o módulo dentro de um prazo determinado. Os desenvolvedores da melhor solução ganham um prêmio financeiro (Birkinshaw and Goddard 2009).

É objetivo deste trabalho analisar os programadores brasileiros quanto à participação (por meio da quantidade de competições), classificação e linguagens de programação favoritas na plataforma TopCoder. Esta análise tem como intuito verificar como esses desenvolvedores do Brasil se comparam a desenvolvedores de outros países. Desta forma, pretende-se entender o potencial das atividades de desenvolvimento de *software* utilizando *crowdsourcing* no Brasil. Além disso, os resultados desta pesquisa poderiam ser utilizados por outras empresas interessadas em explorar o potencial de atividades colaborativas como o *crowdsourcing* em seus contextos. Vale ressaltar que incentivar a participação brasileira em atividades de *crowdsourcing* pode criar novas oportunidades de renda e mobilidade social (Kittur, 2013).

O restante deste artigo está organizado como segue. A próxima seção apresenta o conceito de *crowdsourcing* e detalha a plataforma TopCoder. A seção 3 descreve a metodologia utilizada neste trabalho, incluindo como os dados foram coletados. As próximas seções, 4 e 5, apresentam os resultados do trabalho bem como a discussão dos mesmos. Finalmente, a seção 6 descreve as conclusões e trabalhos futuros.

## **2. *Crowdsourcing* e o TOPCODER**

### **2.1. Visão Geral**

O termo *crowdsourcing* descreve um novo modelo de negócio baseado na *web* que aproveita as soluções de uma rede distribuída de pessoas por meio de um “convite” aberto à apresentação de propostas. Uma das importâncias do *crowdsourcing* é que ele permite o estabelecimento de espaços de idéias compartilhadas para a produção de inovação. Além disso, o *crowdsourcing* pode promover maneiras de como o trabalho deve ser feito, de maneira rápida, sob controle e baixo custo (Malone et al. 2011).

Um exemplo bastante conhecido de *crowdsourcing* é o *Amazon Mechanical Turk*<sup>1</sup> (AMT) da empresa Amazon. Este serviço utiliza a inteligência humana para realizar tarefas chamadas de *Human Intelligence Tasks*. Este serviço permite que empresas acessem programaticamente uma mão-de-obra diversificada sob demanda. A justificativa por trás desta plataforma é que os humanos conseguem ser mais eficientes do que os computadores em algumas tarefas como por exemplo, identificar objetos em uma imagem ou vídeo, transcrever áudios, remover arquivos duplicados a partir de um diretório, entre outros exemplos. O AMT disponibiliza APIs (*Application Programming Interfaces*) para que os desenvolvedores acessem milhares de trabalhadores remotos. Além disso, as APIs permitem integrar programaticamente os resultados obtidos diretamente em projetos corporativos ou acadêmicos. Por exemplo, Hara e colegas (2013) utilizam o AMT junto com *Google Street View* para que a multidão identifique problemas de acessibilidade em calçadas de cidades dos Estados Unidos.

## 2.2. A utilização de *Crowdsourcing* por Empresas

Existem vários casos de empresas que adotam o *crowdsourcing*, incluindo a Wikipédia (Huberman et al 2009) e a Innocentive que entrega prêmios pelas soluções desenvolvidas (Zatti 2013). Um dos problemas que a Innocentive disponibilizou para a multidão<sup>2</sup> foi o vazamento de óleo no golfo do México. Neste caso, 900 soluções foram propostas para o problema. Outro exemplo de uso do serviço da Innocentive foi da empresa farmacêutica Roche que queria encontrar a melhor forma de medir o volume e a qualidade de espécimes clínicos. Em 2008, a empresa patrocinou um concurso recebendo, depois de dois meses, 113 propostas, isto é, muitas soluções em um curto espaço de tempo (Malone et al 2011).

A Netflix, uma empresa de TV por internet dos Estados Unidos com mais de 37 milhões de assinantes em 40 países<sup>3</sup>, em 2006, criou uma campanha à multidão em geral para encontrar um algoritmo de recomendação mais rápido que o seu. Ela ofereceu US\$1.000.000 para a equipe vencedora (Zatti 2013, Schweissguth 2012). O sucesso foi tão satisfatório que, segundo (Lohr 2009), a empresa pretendia apresentar um novo concurso onde os competidores teriam de dar relevância para dados demográficos e comportamentais com o intuito de modelar os “gostos” dos clientes.

A empresa Dell possui uma plataforma de *crowdsourcing* chamada *IdeaStorm*, onde já conseguiu integrar 280 sugestões e ideias em sua linha de produtos, como a sugestão de montar uma linha de computadores baratos rodando Linux. Seguindo a mesma linha da empresa Dell, o Innovation Jam, definido como um gigantesco brainstorming virtual onde funcionários, clientes e mesmo familiares de funcionários contribuem ativamente, e em tempo real, para a geração de ideias (Schneider et al. 2011).

Além desses, o My Starbucks Idea, é uma plataforma para colher ideias com uma equipe dedicada em colher informações e envolver a comunidade nos seus

---

<sup>1</sup> <http://aws.amazon.com/pt/mturk/>

<sup>2</sup> *Crowd*.

<sup>3</sup> <https://signup.netflix.com/MediaCenter>

problemas. Hoje ela é referência em gestão de ideias no mundo todo (Zatti 2013). Tal iniciativa mostra como o *crowdsourcing* busca a determinação das necessidades dos usuários e o que eles querem, a fim de conhecê-los (Schweissguth 2012).

Finalmente, um último exemplo é da Procter & Gamble, uma empresa pioneira em desenvolver soluções de problemas por meio do *crowdsourcing*: ela já realiza esta atividade há mais de 15 anos (Zatti 2013).

### 2.3. *Crowdsourcing* e Engenharia de Software: A plataforma TopCoder

No contexto de engenharia de *software*, a plataforma TopCoder é um dos exemplos mais famosos de *crowdsourcing*. Esta plataforma administra competições de programação chamadas de SRMs (*Single Round Matches*). Estas competições são realizadas quinzenalmente. Esta plataforma foi criada em Novembro de 2000 e vendida para a empresa Appirio em Setembro de 2013 (Lunden 2013). Ela reúne especialistas do mundo todo em design, desenvolvimento, testes entre outras áreas da engenharia de software. O objetivo do TopCoder é beneficiar tanto os *coders* (programadores de sistemas) e quanto as empresas. Os *coders* tem a oportunidade de serem reconhecidos, de melhorar suas habilidades, demonstrar e receber recompensas por seus conhecimentos. Por outro lado, as empresas buscam boas soluções feitas de maneira eficiente, isto é, rápida e com menor custo.

Com mais de 600.000 coders cadastrados na plataforma, a comunidade TopCoder constrói *softwares* e plataformas para diversos clientes. Isto é acompanhado por testes contínuos e processos de gestão, assim, a plataforma coloca para seus usuários uma nova forma de resolução de problemas que é alcançada através de desenvolvedores espalhados ao redor do mundo.

Na prática, o TopCoder “quebra” os projetos de seus clientes em pedaços e oferece os mesmos para a multidão de desenvolvedores freelances como um ou mais desafios competitivos. Por exemplo, um projeto pode começar com um desafio para resultar na melhor ideia de produto de *software*. Um segundo desafio pode prover um alto nível de detalhamento dos objetivos dos projetos e desafiar desenvolvedores para criar os documentos que melhor traduzem os requisitos dos sistemas<sup>4</sup>. O documento de especificação vencedor pode tornar-se a base do próximo desafio em que outros desenvolvedores competem para o desenho da arquitetura do sistema, especificando as partes necessárias do *software* e as conexões entre eles. Outros desafios são lançados para desenvolver, separadamente, cada pedaço e depois integrá-los. Finalmente, um último desafio pode ser lançado para encontrar e corrigir *bugs* em diversas partes do sistema (Malone et al. 2011).

## 3. Metodologia

Um programa em Java (*parser*) foi desenvolvido para extrair dados de documentos HTML obtidos a partir da plataforma TopCoder. O programa inicia com a obtenção do

---

<sup>4</sup> Existe um fórum na página do TopCoder que permite aos desenvolvedores consultarem os clientes para mais detalhes, e essas perguntas e respostas são visíveis para todos os competidores.

document HTML da página de estatísticas do TopCoder<sup>5</sup>, depois obtém os dados relevantes de cada coder disponíveis no documento como por exemplo: nome, nacionalidade, *rating* (pontuação dentro da plataforma), quantidade de competições, linguagem de programação mais usada e por fim armazena esses dados em um banco de dados, desta forma permitindo consultas e análises. A biblioteca *jsoup*<sup>6</sup> foi utilizada na construção deste *parser* (Heaton 2002).

Considerando que este trabalho visa analisar, quantitativamente, o desempenho dos *coders* brasileiros no TopCoder, mais especificamente em relação aos países que formam o BRICS, o Leste Europeu e a América Latina, foi necessário realizar consultas nos dados obtidos que retornassem resultados numéricos desses países, tais como: (i) número máximo, mínimo, média e desvio padrão de competições dos *coders*; (ii) quantidade de *coders*; (iii) quantidade de linguagem-padrão escolhida; (iv) e classificação.

Os resultados apresentados na próxima seção são relatados a partir dos *coders* ativos na classificação disponível no site da comunidade<sup>7</sup>. Essa classificação representa os *coders* ativos ordenados, de maneira decrescente, por pontuação. Um *coder* é considerado ativo quando participa de pelo menos uma competição no período de 180 dias. Estes, por sua vez, estão distribuídos em 134 países, incluindo o BRICS (Brasil, Rússia, Índia, China e África do Sul), 15 países da América Latina e 14 do Leste Europeu.

## 4. Resultados

### 4.1. Resultados Gerais

A Tabela 1 a seguir apresenta os dez Países mais bem colocados no Topcoder, bem como os melhores e os últimos *coders* colocados por país.

**Tabela 1. Classificação dos *coders* por País**

Colocação	País	Número de <i>coders</i> ativos	Melhor <i>coder</i> colocado	Último <i>coder</i> Colocado
1°	Rússia	531	1°	8.822°
2°	Japão	653	9°	8.874°
3°	China	1432	7°	8.883°
4°	Ucrânia	232	13°	8.866°
5°	Polónia	204	6°	8.886°
6°	Bielorrússia	92	2°	8.725°
7°	Estados Unidos	579	16°	8.905°

<sup>5</sup> <http://community.topcoder.com/tc?module=AlgoRank>

<sup>6</sup> <http://jsoup.org>

<sup>7</sup> <http://community.topcoder.com/tc>

8º	Croácia	78	69º	8.518º
9º	Iran	94	32º	8.844º
10º	Coreia do Sul	122	70º	8.885º

Em relação ao número de competições, até a escrita deste artigo tinham sido realizadas 603 competições do tipo SRM. O *coder* mais ativo da plataforma, isto é, aquele que participou do maior número de competições é um *coder* dos EUA com 401 competições. Entretanto, o número médio de competições por *coder* ativo é 21,81, o desvio padrão é de, aproximadamente, 37,99 e o mínimo é de 1 competição. Deve-se observar que quantidade de participação em competições não implica em melhor classificação, visto que o *coder* mais ativo ocupa a 1.116ª posição.

Em relação às linguagens de programação, os *coders* mais ativos escolhem como linguagem-padrão: C++, Java, C#, Python e Visual Basic (ver Figura 1).

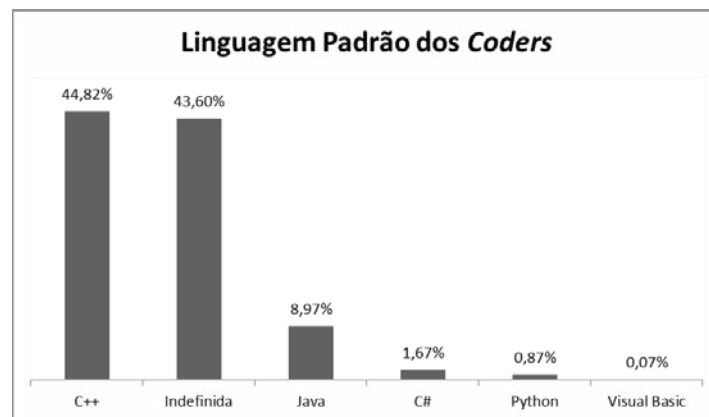


Figura 1. Linguagem de programação padrão escolhida pelos *coders*

#### 4.2. Resultados dos brasileiros

O Brasil está presente no TopCoder e relativamente bem colocado: entre os 13 primeiros Países. Quando da coleta de dados, o Brasil possuía 147 *coders* presentes na classificação geral dos programadores ativos, com o melhor colocado ocupando a 45ª posição (Tabela 2) e o último colocado ocupando a 8825ª posição.

Tabela 2. Classificação dos dez brasileiros mais bem colocados

<i>Coder</i>	Competições	Posição	Linguagens
Brasileiro 1	82	45º	C++
Brasileiro 2	49	151º	C++
Brasileiro 3	92	165º	C++
Brasileiro 4	58	184º	C++
Brasileiro 5	47	292º	C++
Brasileiro 6	291	438º	C++
Brasileiro 7	64	452º	C++
Brasileiro 8	20	549º	C++
Brasileiro 9	38	668º	C++
Brasileiro 10	92	700º	C++

Os brasileiros participaram em média de 25,28 competições, apenas um pouco acima da média mundial. O brasileiro mais participativo esteve presente em 342 competições diferentes. Por outro lado, o *coder* brasileiro que menos participou de competições, esteve presente em apenas 1. Levando em consideração a linguagem de programação mais utilizada pelos brasileiros, o C++ é a linguagem padrão, refletindo a mesma linguagem padrão mundial, conforme a Figura 2.

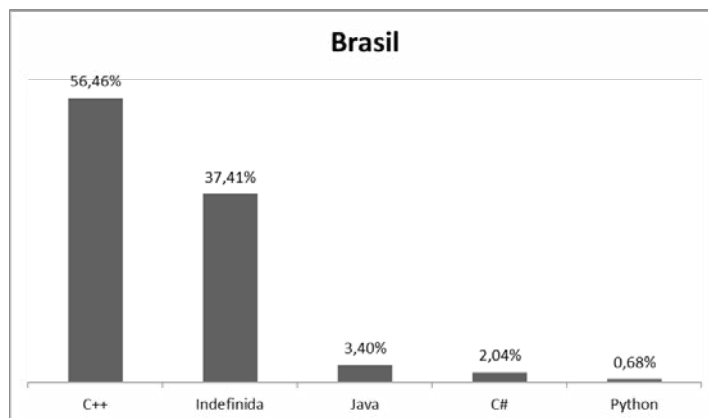


Figura 2. Linguagem de programação padrão escolhida pelos brasileiros

### 4.3. Brasil em relação a América Latina

Para fins de comparação, tendo os dados obtidos neste trabalho como base, definiu-se que seria traçado, primeiramente, um paralelo entre os programadores brasileiros e os pertencentes à América Latina. Como se pode observar, o Brasil possui o maior número de *coders* ativos, bem como o melhor colocado no *ranking* da TopCoder. Dentre os países da América Latina, o Brasil é o mais bem colocado. O país que mais se aproxima no referido *ranking* é a Argentina, ocupando a 29ª posição, seguida pelo Peru em 33°.

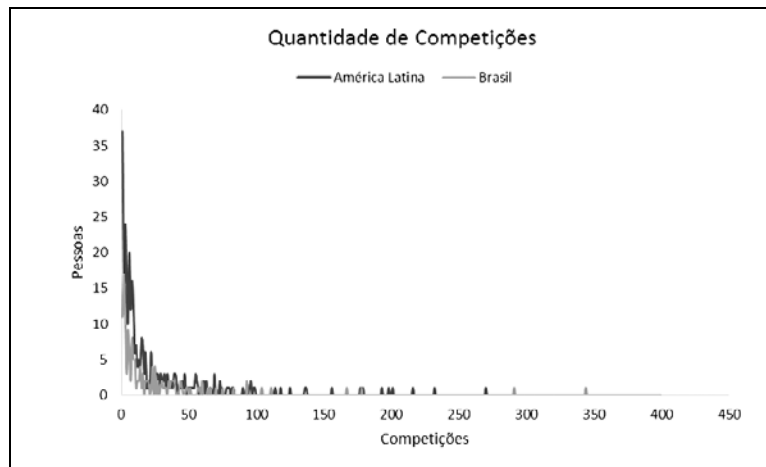
Tabela 3. Classificação dos *coders* por Países da América Latina

Colocação	País	Número de <i>coders</i> ativos	Melhor <i>coder</i> colocado	Último <i>coder</i> colocado
13°	Brasil	147	45°	8825°
29°	Argentina	44	274°	8292°
33°	Peru	86	360°	8898°
38°	Mexico	49	478°	8765°
44°	Colômbia	50	444°	8760°
48°	Bolivia	44	475°	8807°
54°	Venezuela	18	1403°	8871°
58°	Republica Dominicana	12	920°	8074°

Além disso, os dados de número de competições e linguagem de programação foram analisados. No que diz respeito ao número de competições a média entre o Brasil

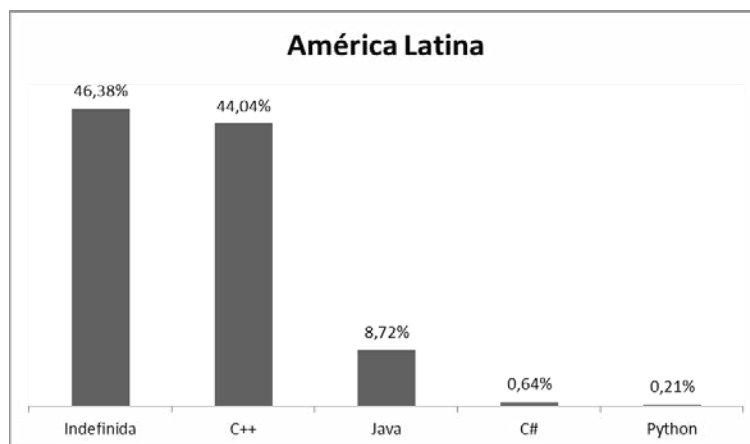


e a América Latina é a praticamente a mesma por desenvolvedor ranqueado, como pode ser visto na Figura 3.



**Figura 3. Comparação de quantidades competições dos coders brasileiros em relação a América Latina.**

Quanto à linguagem de programação, os brasileiros, bem como o restante dos Países latino-americanos, utilizam mais o C++ como linguagem padrão, seguido pelo Java (Figura 4). Este padrão é similar ao do resto do mundo conforme apresentado na Figura 1 anteriormente.



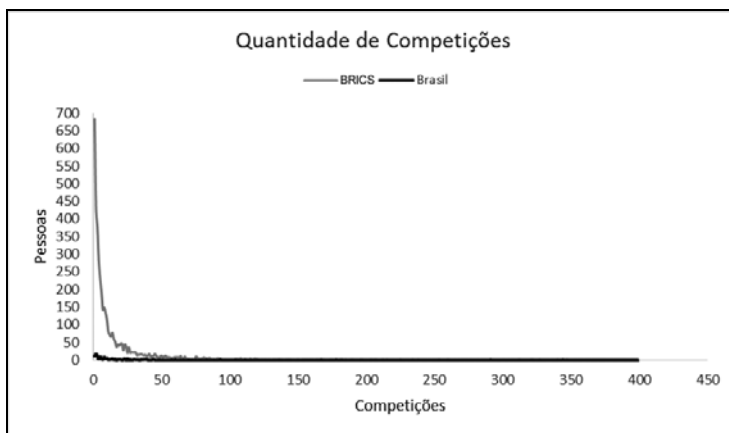
**Figura 4. Linguagem de programação padrão escolhida pelos coders da América Latina.**

#### 4.4. Brasil em relação ao BRICS

É interessante analisar também como o Brasil está em relação aos países que formam o BRICS, pois os brasileiros fazem parte deste grupo político. A partir dos dados foram analisadas as seguintes características: número de competições, *ranking* e linguagem de programação. Em relação ao número de competições, a média do Brasil é maior comparado a do BRICS que é de 16,74 competições por *coder* ativo. Porém, a Figura 5 mostra que os *coders* pertencentes aos países que formam as siglas BRICS participam de um maior número de competições que os brasileiros. Já o número máximo de



competições, o brasileiro que mais competiu perde para o russo (primeiro do *ranking* geral) que competiu 346 vezes.



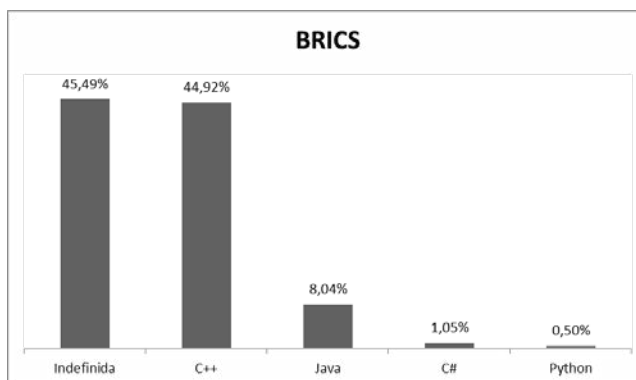
**Figura 5. Comparação de quantidades de competições dos coders brasileiros em relação aos outros Países do BRICS.**

Conforme a Tabela 4, o Brasil, em relação aos países do BRICS, é o terceiro colocado e está atrás da China e Rússia. É importante destacar que a Rússia é o país que ocupa a primeira posição e a China a terceira posição da classificação geral dos países. A tabela abaixo mostra a classificação dos países do BRICS em relação a todos os competidores ativos do TopCoder.

**Tabela 4. Classificação dos coders por Países do BRICS**

Colocação	País	Número de coders ativos	Melhor <i>coder</i> colocado	Último <i>coder</i> colocado
1º	Rússia	531	1º	8822º
3º	China	1432	7º	8883º
13º	Brasil	147	45º	8825º
18º	Índia	2264	217º	8906º
36º	África do Sul	11	4º	8767º

Observando a Figura 6, a maioria dos *coders* pertencentes ao BRICS escolhem o C++ como linguagem padrão.



**Figura 6. Linguagem de programação padrão escolhida pelos coders do BRICS**

#### 4.5. Brasil em relação ao Leste Europeu

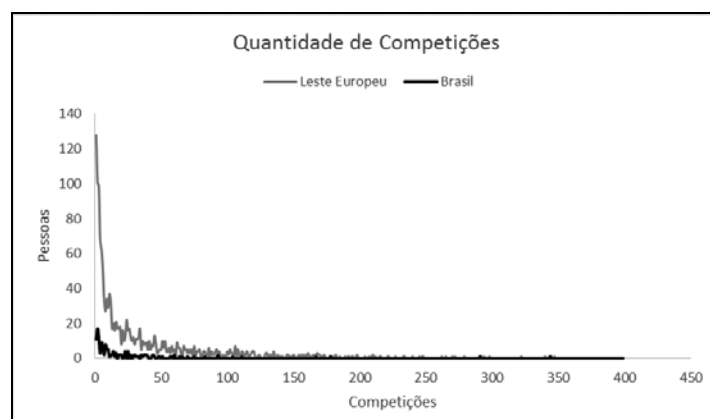
Conforme visto anteriormente, o Brasil está atualmente na 13ª posição da classificação geral. Entretanto, em uma análise comparativa apenas com os países do Leste Europeu, o Brasil se encontra na 6ª colocação se comparado apenas aos 14 países presentes no *ranking* do TopCoder e que fazem parte do Leste Europeu: Belarus, Bulgária, República Checa, Croácia, Eslováquia, Hungria, Letônia, Lituânia, Macedônia, Polônia, Romênia, Rússia, Servia e Ucrânia.

**Tabela 5. Classificação dos dez *coders* mais bem colocados por Países do Leste Europeu.**

Colocação	País	Número de <i>coders</i> ativos	Melhor <i>coder</i> colocado	Último <i>coder</i> colocado
1º	Rússia	531	1º	8822º
4º	Ucrânia	232	14º	8866º
5º	Polônia	204	6º	8886º
6º	Bielorrússia	92	2º	8725º
8º	Croácia	78	69º	8518º
<b>13º</b>	<b>Brasil</b>	<b>147</b>	<b>45º</b>	<b>8825º</b>
14º	Bulgária	69	128º	8897º
17º	Eslováquia	43	60º	7548º
21º	România	106	208º	8538º
28º	Lituânia	12	292º	6518º
30º	República Checa	22	78º	8003º

Outra comparação refere-se a média de competições por desenvolvedor brasileiro que é de 25,285, enquanto que para média do conjunto de Países que compõem o Leste Europeu é de 36,097 competições. Essa disparidade de participação média em competições é comprovada na Figura 7.

A Figura 8 apresenta a escolha das linguagens de programação dos *coders* do Leste Europeu. Novamente, a linguagem C++ apresenta-se como a linguagem mais utilizada.



**Figura 7. Comparação de quantidades de competições dos *coders* brasileiros em relação aos países do Leste Europeu.**

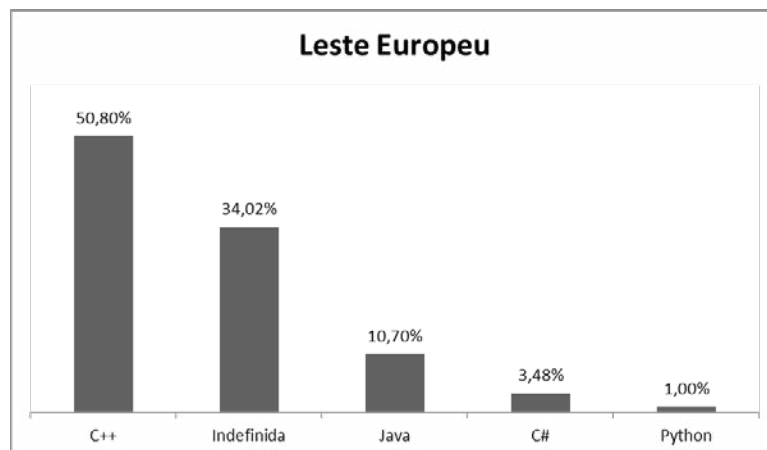


Figura 8. Linguagem de programação padrão escolhida pelos *coders* do Leste Europeu.

## 5. Discussão

### 5.1. Visão Geral

A Figura 9 abaixo apresenta o total de pesquisas feitas, em relação ao número total de pesquisas realizadas no *Google* ao longo do tempo para a palavra “*Crowdsourcing*”. Estas são as pesquisas feitas em todo o mundo no período de outubro de 2009 a maio de 2015. O figura abaixo é resultado de uma pesquisa utilizando o *Google Trends*<sup>8</sup>.

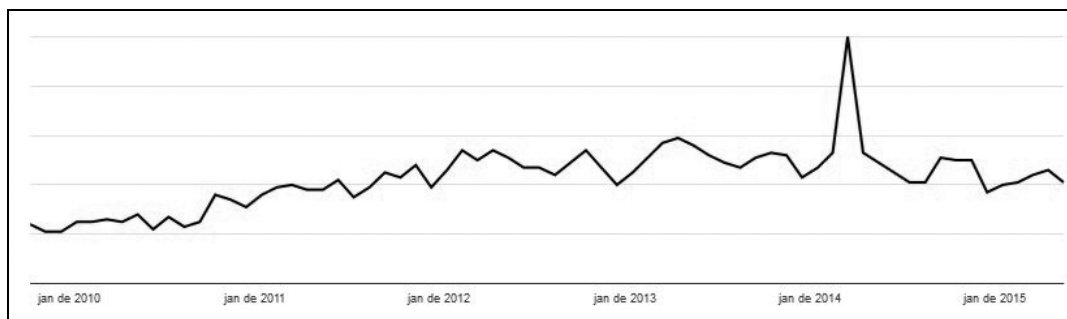


Figura 9. Pesquisa pelo termo *Crowdsourcing* no *Google*

A Figura 9 sugere um crescimento genérico no interesse deste novo modelo de produção. Por ser um tema relativamente recente, nos últimos anos a comunidade científica também tem se esforçado para entender melhor este fenômeno com diversos artigos nas conferências nacionais (SBSC) e internacionais (CSCW e CHI) incluindo (Chan et al, 2016) (Antonio et al, 2014) (Lykourantzou et al, 2016).

Os resultados anteriores sugerem que o Brasil está relativamente bem colocado na plataforma TopCoder, especialmente quando comparado aos outros países da América Latina, mas nem tanto quanto comparados aos outros países do BRICS e os países do Leste Europeu. Entretanto, vale ressaltar que o número de brasileiros ativos no

<sup>8</sup> <http://www.google.com/trends>

TopCoder ainda é pequeno, com apenas 147 desenvolvedores quando da coleta dos dados utilizados neste artigo.

Assim, um estudo empírico sobre a participação brasileira no TopCoder pode trazer incentivos para a indústria nacional, possibilitando investigar fatores para melhoria da participação no TopCoder como mostram as evidências de outros autores (Li et al, 2013). Os autores acreditam que este trabalho contribuirá para o aumento do conhecimento sobre o *crowdsourcing* no contexto do desenvolvimento de *software* pela comunidade científica brasileira. Isto tem o potencial de estimular outros estudos sobre a participação brasileira neste modelo de negócio. Isto é particularmente importante no atual contexto econômico do Brasil, pois o *crowdsourcing* também cria novas oportunidades de renda e mobilidade social em regiões do mundo onde as economias locais podem ser estagnadas ou sem investimento (Kittur, 2013).

## 5.2. Trabalhos Relacionados

Diversos aspectos têm sido estudados no contexto de *crowdsourcing*, e ainda mais recentemente, como o *crowdsourcing* pode trazer ganhos para as atividades de desenvolvimento de *software*. Neste contexto, o *crowdsourcing* pode ser definido como “uma realização de determinadas tarefas de desenvolvimento de *software* em nome de uma organização por um grande grupo indefinido de pessoas externas com os conhecimentos técnicos especializados através de um chamado aberto” (Stol, 2014). Além disso, essa abordagem usa uma multidão online para terceirizar tarefas, incluindo requisitos, projeto, codificação, testes, evolução e documentação (Leicht, 2015).

Conforme mencionado anteriormente, a plataforma TopCoder é uma das maiores sendo utilizada no contexto de desenvolvimento de software. Assim, ela tem sido estudada por diferentes pesquisadores. Por exemplo, Douglas (2014) discute algumas características da comunidade TopCoder coletadas através de um *survey* com 2.100 respostas de membros da comunidade. Um dos resultados é que quase a metade dos entrevistados disseram que sua linguagem de desenvolvimento preferida é, não é surpreendentemente, C++ ou Java. Este resultado corrobora com o que foi apresentado neste trabalho.

Outro estudo que analisa a comunidade no TopCoder foi realizado por (Boudreau et al 2014). Estes autores buscaram analisar os fatores que levam à um maior ou menor esforço por parte dos coders. Para tanto, eles realizaram um experimento através de um conjunto de 260 coders escolhidos aleatoriamente para um problema complexo ao longo de 10 dias. Como mencionado antes, a plataforma TopCoder fornece uma infraestrutura pronta para acolher os desafios de programação e tem um sistema de classificação e avaliação de habilidades bem desenvolvidas, que permite a identificação da capacidade de qualquer competidor. Assim, os autores foram capazes de tomar medidas detalhadas e objetivas de esforço, comunicação e performance dos *coders*. Os resultados sugerem que os níveis de esforço são movidos por incentivos financeiros e da interação de outros companheiros de equipe.

De maneira similar a este trabalho, Mattos e colegas (2013) apresentam uma análise de como as empresas brasileiras estão usando as plataformas de colaboração para a inovação aberta. O objetivo do trabalho era identificar as práticas de empresas

quanto a condução de iniciativas de *crowdsourcing*. Os resultados preliminares da pesquisa indicam que os projetos de *crowdsourcing* têm sido bem sucedidos.

Um outro estudo empírico baseado no Topcoder é apresentado por Li e colegas (2013). Com base na literatura existente estes autores propõe 23 fatores de qualidade de software em 2 aspectos: plataforma e projeto. Para tal, eles utilizam um programa rastreador (crawler) para coletar informações de 836 projetos no período de 2003 a 2012. Além disso, eles utilizam um modelo de regressão múltipla para analisar o impacto dos fatores na qualidade de software e sugerem 6 fatores chaves para melhoria da qualidade: pontos de qualidade na plataforma, quantidade de projetos ativos, quantidade dos membros registrados no projeto, pontuação máxima dos desenvolvedores, pontuação do projeto e tamanho da documentação disponível. Além disso, Li e colegas também sugerem 4 aspectos para ajudar na melhoria da qualidade de *software* desenvolvido via *crowdsourcing*: escolher o melhor período para postar o projeto, reduzir a escala do projeto, atrair desenvolvedores mais habilidosos e, finalmente, melhorar a pontuação do projeto. Conforme apresentado, pesquisas em Crowdsourcing vêm crescendo ao longo do tempo, e este trabalho teve como motivação contribuir com a área. Além disso, os dados coletados nesta pesquisa estão sendo armazenados para que se possa em um trabalho futuro analisar o histórico de participação de coders brasileiros, bem como de outros países presentes na plataforma.

## 6. Conclusão

Essa pesquisa buscou contribuir para o conhecimento sobre o *crowdsourcing* no contexto de desenvolvimento de *software* através de um estudo da plataforma TopCoder. Além disto, este artigo visa estimular estudos adicionais sobre a participação brasileira nesta nova forma de trabalho, além de permitir um maior entendimento sobre o potencial das atividades de desenvolvimento de *software* utilizando *crowdsourcing* no Brasil. Foram coletados dados sobre classificação de coders e países a partir das páginas HTML da própria plataforma, utilizando um *parser* de documentos HTML. Após análises, concluímos que o Brasil está relativamente bem colocado no geral, bem como entre todas as comparações realizadas nesta pesquisa, ou seja, o Brasil em relação à América Latina, ao BRICS e Leste Europeu. Entretanto, vale ressaltar que o número de brasileiros ativos no TopCoder ainda é pequeno, com apenas 147 desenvolvedores. A análise contínua destes dados por um longo espaço de tempo poderá revelar a evolução quanto ao aumento da participação por países, trabalho este que está em andamento.

Vale ressaltar uma última relevância deste trabalho para comunidade de sistemas colaborativos: os resultados apresentados buscaram estimular uma contribuição maior em pesquisas empíricas no contexto do crowdsourcing ao invés de focar na construção de plataformas e aplicações. Em outras palavras, entender a participação da multidão (*crowd*), ou pelo menos parte dela, em uma plataforma de crowdsourcing pode servir de ponto de partida para a investigação de fatores que visem o aumento da participação no Topcoder (Li e colegas, 2013). Além disso, plataformas de crowdsourcing também estão relacionadas a comunidades online, logo o entendimento da adoção e retenção destas comunidades é um importante aspecto de pesquisa em Sistemas Colaborativos.

## Agradecimentos

Esta pesquisa é financiada pelo CNPq através dos processos 406692/2013-0, 440880/2013-0 e 310468/2014-0.

## 7. Referências

- Alves, L., et al. (2011) “Um aplicativo baseado em inteligência coletiva para compartilhamento de rotas em redes sociais”. VIII Simpósio Brasileiro de Sistemas Colaborativos.
- Antonio M., et al. (2014) “Filtragem colaborativa de rotas de ônibus usando dados sobre a acessibilidade das vias urbanas”. Anais do XI Simpósio Brasileiro de Sistemas Colaborativos. Curitiba, PR, Brasil.
- Birkinshaw, J. and Goddard, J. (2009) “Qual é seu modelo de gestão?”. HSM Management 75.
- Boudreau, K., et al. (2014) “From Crowds to Collaborators: Initiating Effort & Catalyzing Interactions Among Online Creative Workers”. Social Science Research Network. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2384068](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2384068).
- Brabham, D. (2008) “Crowdsourcing as a model for problem solving”. The International Journal of Research into New Media Technologies. Vol. 14.
- Chan J., et al. (2016) “Improving Crowd Innovation with Expert Facilitation”. Proceedings of the 19th CSCW Conference. San Francisco, CA, USA.
- Douglas, J. (2014) “Getting to Know You!”. <http://www.topcoder.com/blog/results-of-getting-to-know-you-survey/>
- Hara K., et al. (2013) “Combining Crowdsourcing and Google Street View to Identify Street Level Accessibility Problems”. CHI'13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Heaton, J. (2002) “Programming Spiders, bots and Aggregations in Java”. San Francisco, USA.
- Huberman, B., et al. (2009) “Crowdsourcing, attention and productivity”. Journal of Information Science. pp. 758-765.
- Kittur, A., et al. (2013) “The Future of Crowd Work”. In Proceedings of the 2013 conference on Computer supported cooperative work (CSCW 2013). USA, San Antonio, 1301-1318.
- Leicht, N., et al. (2015) “Crowdsourcing in Software Development: A State-of-the-Art Analysis”. 28th Bled eConference. Bled, Slovenia.
- Li, K., et al. (2013) “Analysis of the key factors for software quality in crowd sourcing development: An empirical study on TopCoder.com”. Proceedings of the 2013 IEEE 37th Annual Computer Software and Applications Conference, pp. 812-817.
- Lohr, S. (2009) “Netflix Awards \$1 Million Prize and Starts a New Contest”. <http://bits.blogs.nytimes.com/2009/09/21/netflix-awards-1-million-prize-and-starts-a-new-contest/>.

- Lunden, I. (2013) “Appirio Buys TopCoder To Add More Crowdsourcing, And 500K Developers And Designers, To Its CloudSpokes Network”. <http://techcrunch.com/2013/09/17/appirio-buys-topcoder-to-add-more-crowdsourcing-and-500k-developers-and-designers-to-its-cloudspokes-network/>
- Lykourantzou, I., et al. (2016) “Team Dating: A Self-Organized Team Formation Strategy for Collaborative Crowdsourcing”. Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing System. San Jose, CA, USA.
- Malone, T., et al. (2010) “The collective intelligence genome”. MIT Sloan Management Review. Vol. 41.
- Malone, T., et al. (2011) “The big idea. The age of hyperspecialization”. Harvard Business Review.
- Mattos, C., et al. (2013) “The Use Of Collaborative Platforms For Improved Open Innovation: Brazilian Context”. International Conference on Industrial Engineering and Operations Management. [http://www.abepro.org.br/biblioteca/ICIEOM2013\\_STO\\_170\\_979\\_21786.pdf](http://www.abepro.org.br/biblioteca/ICIEOM2013_STO_170_979_21786.pdf)
- Schneider, D., et al. (2011) “Multidões: a nova onda do CSCW?”. VIII SBSC.
- Schweissguth, S. (2012) “Learning From the Crowdsourcing Efforts at Netflix”. <http://crowdsourc.com>.
- Stol, K. and Fitzgerald, B. (2014) “Two's company, three's a crowd: a case study of crowdsourcing software development”. ICSE.
- Zatti, R. (2013) “Crowdsourcing e as peças de uma nova forma de fazer as coisas”. <http://www.slideshare.net/rafazatti/e-book-crowdsourcing-e-as-peas-de-uma-nova-forma-de-fazer-as-coisas>.