Pessoas que talvez você conheça

Recomendação de amigos utilizando MapReduce via Hadoop

Michel Francisco Nagme Rachid Costa – RA: 082303
Ricardo Amaral de Pretto – RA: 072266

Motivação

"Se duas pessoas em uma rede social tem um amigo em comum, então existe uma maior probabilidade de que elas próprias se tornem amigas em algum ponto no futuro."

- *David Easley e Jon Kleinberg,*

em Networks, Crowds, and Markets: Reasoning About a Highly Connected World

Problema

1,3 bilhão de usuários ativos

540 milhões de usuários ativos

300 milhões de usuários ativos

270 milhões de usuários ativos

Algoritmo MapReduce

Rodrigo	João, Juliana, Carlos
João	Rodrigo, Carlos, Igor, Amanda
Juliana	Rodrigo, Igor
Carlos	Rodrigo, João, Igor
Igor	João, Juliana, Carlos
Amanda	João, Beatriz
Beatriz	Amanda

Map Phase

1. Para cada usuário “A” e seus amigos “An”, gerar mapas **<key, (value1, value2)> = <A, (An,*)>**, onde * indica ao algoritmo que A e An já são amigos.

2. Sejam “A1” e “A2” amigos de “A”, gerar mapa **<key, (value1, value2)> = <A1, (A2, A)>**. Fazer isso para todas as combinações dos “An” amigos de “A”.

Map Phase

1. Para cada usuário “A” e seus amigos “An”, gerar mapas **<key, (value1, value2)> = <A, (An,*)>**, onde * indica ao algoritmo que A e An já são amigos.

Rodrigo **João, Juliana, Carlos**

<Rodrigo, (João, *)>
<Rodrigo, (Juliana, *)>
<Rodrigo, (Carlos, *)>

Map Phase

1. Sejam “A1” e “A2” amigos de “A”, gerar mapa **<key, (value1, value2)> = <A1, (A2, A)>**. Fazer isso para todas as combinações dos “An” amigos de “A”.

Rodrigo **João, Juliana, Carlos**

<João, (Juliana, Rodrigo)> **<Juliana, (João, Rodrigo)>**
<João, (Carlos, Rodrigo)> **<Juliana, (Carlos, Rodrigo)>**

<Carlos, (João, Rodrigo)>
<Carlos, (Juliana, Rodrigo)>

Reduce Phase

1. Para cada combinação de mapas com chave “A”, contar quantas vezes outro usuário “B” qualquer foi sugerido por um amigo mútuo. Se qualquer um dos amigos mútuos estiver indicado por *, então descartar a sugestão (pois A e B já são amigos).

2. Ordenar os resultados de forma decrescente de acordo com o número de amigos em comum.

Reduce Phase

1. Para cada combinação de mapas com chave "A", contar quantas vezes outro usuário "B" qualquer foi sugerido por um amigo mútuo. Se qualquer um dos amigos mútuos estiver indicado por *, então descartar a sugestão (pois A e B já são amigos).

<Rodrigo, (João, *)>
<Rodrigo, (João, Carlos)>
=
Rodrigo e João já são amigos

<Rodrigo, (Igor, João)>
<Rodrigo, (Igor, Juliana)>
<Rodrigo, (Igor, Carlos)>
=
Igor(3) é sugerido como amigo para Rodrigo

Resultado

Rodrigo	Igor(3:[Juliana,João,Carlos]), Amanda(1:[João])
João	Juliana(2:[Igor,Rodrigo]), Beatriz(1:[Amanda])
Juliana	João(2:[Igor,Rodrigo]), Carlos(2:[Igor,Rodrigo])
Carlos	Juliana(2:[Rodrigo,Igor]),Amanda(1:[João])
Igor	Rodrigo(3:[João,Carlos,Juliana]), Amanda(1:[João])
Amanda	Rodrigo(1:[João]), Carlos(1:[João]), Igor(1:[João])
Beatriz	João(1:[Amanda])

Implementação

0	1,2,3
1	0,3,4,5
2	0,4
3	0,1,4
4	1,2,3
5	1,6
6	5

0	4 (3:[2,1,3]), 5 (1:[1])
1	2 (2:[4,0]), 6 (1:[5])
2	1 (2:[4,0]), 3 (2:[4,0])
3	2 (2:[0,4]), 5 (1:[1])
4	0 (3:[1,3,2]), 5 (1:[1])
5	0 (1:[1]),3 (1:[1]),4 (1:[1])
6	1 (1:[5])

Obrigado!

- DB Tsai. "H*adoop M/R to Implement "People You Might Know" Friendship Recommendation"*. 22 Jan. 2013. Web. 23 Set. 2014. <http://dbtsai.github.io/blog/2013/hadoop-mr-to-implement-people-you-might-know-friendship-recommendation/>.

- Matthews, Paul. "Teaching MapReduce." *Paul Matthews Blog @ UWE*. 4 Mar. 2014. Web. 23 Set. 2014. <http://www.cems.uwe.ac.uk/~pmatthew/blog/2014/03/04/teaching-mapreduce>.

- ""People You May Know" Friendship Recommendation with Hadoop." *Important Fish*. 14 Julho 2013. Web. 23 Set. 2014. <http://importantfish.com/people-you-may-know-friendship-recommendation-with-hadoop/>.