Report of ASC Student Supercomputer Challenge 2017 Preliminary Contest

Zhibin Xie, Zhen'ning Wang, Jiafang Xue, Guohua Tian
School of Information Engineering,
NanChang University

Xianling Dong
School of Science,
NanChang University

***Abstract*— NSBSS.**

1. Introduction

We are a group of youths loving computer science who have a huge dream, and working hard for reaching higher. We gathered together due to strong interest in supercomputer.

NSBSS
March 5, 2017

1.1. Team's Information

NSBSS Team is ASC Team of NanChang University...

1.2. Other Informations

etc.

2. HPL Test

1) HPL install

In this test and optimize, we chose Intel Optimized LINPACK.

2) Test Step

2.1) PxQ

Because Intel Optimized LINPACK use all CPU core in multi-thread by default, so we set P x Q to 1 x 1.

2.2) BLAS
There are many blas, such as openblas, mkl, gogoblas and atlas. Because we are using intel xeon phi, MKL will be the best choice

2.3) N
Experience tells us, N can be calculated by

$$N^2 * 8 = MEM * 80\%$$

so N should have been 82900, but in our test, when N is set to 35000, GFlops reach to 1.7GFlops, which is much higher when N is 829000

N	33000	34000	35000	36000	37000
GFlops	1702	1701.48	1714.1	1708.11	1501.21
Time (S)	14.08	15.4	16.68	18.21	22.5

2.4) NB

NB	232	256	296
GFlops	1709.86	1714.1	1702.39

2.1. Best Score

Finally, we reached at 1.7Gflops.

2.2. Problems

Problems we had meet here.

2.3. etc.

Subsection text here.

3. MASNUM_WAVE

Details of MASNUM_WAVE, 2 line at least....
line.2

3.1. Information I

Subsection text here.

3.2. Information II

Subsection text here.

4. Deep learning contest

In-depth learning is a learning algorithm that simulates the cognitive pattern of the human brain, and the current Baidu Research Institute uses the strategic direction of data mining and artificial intelligence as an important development strategy of the 21st century.

PaddlePaddle is originated in Baidu's open source depth learning platform.And it's easy to use.

4.1. Analysis of the problem

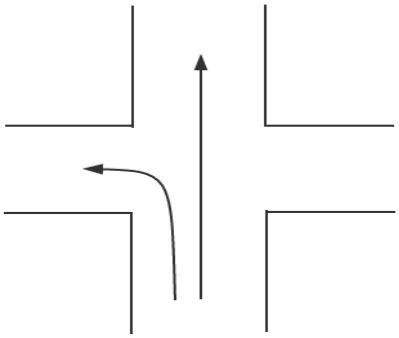
Data are sampled from mobile phone APP.Using the dataset which are aggregated about 50 days from 00:00 a.m. on March 1st to 08:00 on April 20th for training. Your task is to predict each link's speed for every 5 minutes from 08:00 to 10:00 on April 20th.

The model outputs are evaluated by the Root Mean Squared Error (RMSE) of predictions.

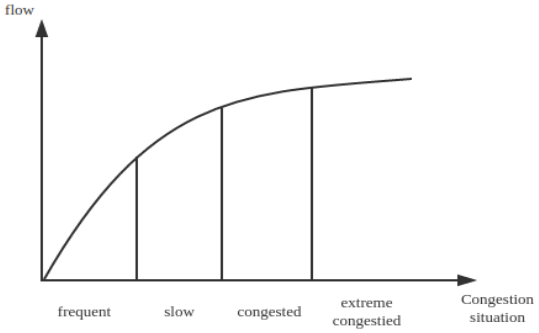
$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_{actual,i} - X_{model,i})^2}{n}}$$

[1] $X_{actual,i}$ is actual value of moment i, while $X_{model,i}$,i is your prediction moment of i

By simply observing the node graphs, we map the traffic nodes, for example, simple crossroads



Corresponding to the four traffic states.



The result of clustering is the four eigenvectors (1, 2, 3, 4) corresponding to the state defined by the topic. The four eigenvectors represent smooth, slow, crowded and very crowded.

4.2. Questions

After we downloaded the data, we found that the amount of data is large, some of the data there is a large interference, so the data processing is very important.We also have little understanding of the depth of learning. so it is difficult to find theoretical support. And the results of the prediction, and the calculation is unclear.

4.3. Preliminary solution

4.3.1. Data processing. Through the analysis of the title and data, we selected the daily time from 00:00-02:00, 04:00-06:00, 08:00-10:00, 12:00-14:00, 16:00-18:00 and 20:00-22:00 six time periods as the representative.In order to reduce the error, we take the average of the time to process the data.

4.3.2. Algorithm selection. Algorithm we used RMSProp, RMSProp is a very efficient algorithm, RMSProp slightly improved AdaGrad, making the algorithm no longer as radical as AdaGrad it is moving average of squared gradients.

The equations of this method as follows:

$$\nu(w,t) = \rho\nu(w,t-1) + (1-\rho)(\nabla Q_i(w))$$

$$w = w - \frac{\eta}{\sqrt{\nu(w,t)+\varepsilon}}(\nabla Q_i(w))$$

4.3.3. Result analysis. We perform the error analysis by plotting the original data, and comparing the resulting data

Relative error:

$$rerr = \frac{X_{actual,i} - X_{model,i}}{X_{actual,i}}$$

Absolute relative error:

$$mrerr = \frac{1}{N}\sum_{1}^{N}\left(\frac{X_{actual,i} - X_{model,i}}{X_{actual,i}}\right)$$

4.4. Algorithm improvement

Adam is a recently proposed algorithm that is similar to the RMSprop comparison with the momentum. The procedure is similar to:

$$m(w,t) = \beta_1 m(w,t-1) + (1-\beta_1)\nabla Q_i(w)$$

$$v(w,t) = \beta_2 v(w,t-1) + (1-\beta_2)(\nabla Q_i(w))^2$$

$$w = w - \frac{\eta}{\sqrt{\nu(w,t)+\varepsilon}}$$

And we set eps=1e-6, bata1=0.9, beta2=0.999.

1. A

4.5. Parameter configuration

Choose a suitable learning rate is very difficult, in the beginning, we set the way through the schedule learning learning method, that is pre-designed a certain series of iterations to increase or increase learning rate.As the learning rate increases, we find that the relative error begins to increase.Finally we set the learning rate to 1e-6.

5. Conclusion

No conclusion yet...

6. Appendix

No appendix yet.

6.1. A.

Appendix A.

References

[1] No reference yet...