Reservoir Sampling

Anže Pečar, Miha Zidar

***Abstract*—The abstract goes here.**

***Index Terms*—sampling, reservoir, machine learning.**

I. Introduction

THE problem with random sampling is to select a random sample of size n from a set of size N. Many algorithms have been developed for this problem when the value of N is known beforehand. Some problems, that we encounter in the real world, do not have a specified N or N cannot be determined efficiently. Those kind of problems will be the focus of this paper. We shall take a look at older algorithms such as Algorith R which was developed to efficiently and accurately create a random sample from a tape in one pass, as well as newer algorithms, such as Open learning[TODO]

Many reservoir algorithms make the assumption that the data does not change over time. Alogrithm R, for an example, samples starting data with a greater frequency than the data at the end of the tape[citation needed], which is efficient but it fails to react to emarging trends in the data. If we are interested in catching trends in our data a different kind of algorithm needs to be found. [TODO]

December 5, 2011

II. Algorithm R and its improvements

A. Motivation

In order to better understand the complex algorithms in use today we first need to understand the idea behind algorithms used in the past. In this section we shall describe Algorithm R and some of its improvement Algorithm Z.

B. Algorithm R

[TODO]

C. Algorithm Z

[TODO]

III. Open learning

[TODO]

IV. Conclusion

The conclusion goes here.

Appendix A
Appendix title

Appendix one text goes here.

Appendix B

Appendix two text goes here.

Acknowledgment

The authors would like to thank...

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

Anže Pečar Biography text here.

PLACE
PHOTO
HERE

Miha Zidar Biography text here.