Reward Maximization for Task allocation in Opportunistic crowdsensing networks

Michael Shell, *Member, IEEE,* John Doe, *Fellow, OSA,* and Jane Doe, *Life Fellow, IEEE*

Abstract—The abstract goes here.

Index Terms—Computer Society, IEEEtran, journal, LaTeX, paper, template.

1 Introduction

THIS demo file is intended to serve as a "starter file" for IEEE Computer Society journal papers produced under LaTeX using IEEEtran.cls version 1.8a and later. I wish you the best of success.

mds
September 17, 2014

1.1 Subsection Heading Here

Subsection text here.

1.1.1 Subsubsection Heading Here

Subsubsection text here.

2 Conclusion

The conclusion goes here.

Appendix A
Proof of the First Zonklar Equation

Appendix one text goes here.

Appendix B

Appendix two text goes here.

Acknowledgments

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.

PLACE
PHOTO
HERE

Michael Shell Biography text here.

John Doe Biography text here.

Jane Doe Biography text here.

Appendix C
MAB algorithm

C.1 ϵ-greedy-TA

- *M. Shell is with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332. E-mail: see http://www.michaelshell.org/contact.html*
- *J. Doe and J. Doe are with Anonymous University.*

Manuscript received April 19, 2005; revised September 17, 2014.

Algorithm 1 ε-Greedy-TA

Input: K: the number of arms;
R: reward function;
T: the round of times;
ϵ: probability

Output: r:the accumulate reward.

$r = 0$;
$\forall i = 1, 2, ..., K : Q(i) = 0, count(i) = 0$;
for $t = 1, 2, ..., T$ **do**
  if $rand() < \epsilon$ **then**
    randomly choose one from 1,2,...,K and assignment to k
  else
    $k = \arg\max_i Q(i)$
  $v = R(k)$;
  $r = r + v$;
  $Q(k) = \frac{Q(k)*count(k)+v}{count(k)+1}$;
  $count(k) = count(k) + 1$;

if the reward of the arm is very uncertainty, for example, its probability distrubutions contains too much value, then it will need more exploration, and then it need big ϵ value.Vice versa.

C.2 softmax-TA

In softmax algorithm, it acccording to the known mean reward of the arms to decide to choose which arm. If the mean reward is bigger, and the probability of choose this arm is bigger.

And the probability distribution of choose a arm is Boltzmann distributions:

$$P(k) = \frac{e^{\frac{Q(k)}{\tau}}}{\sum_{i=1}^{K} e^{\frac{Q(i)}{\tau}}} \tag{1}$$

Q(i) represent the mean reward upon now; $\tau > 0$ is called "temperature", and when τ is very small, the mean reward bigger one will have higher probability to be choose.

Algorithm 2 softmax-TA

Input: K: the number of arms;
R: reward function;
T: the round of times;
τ: temperature parameter

Output: r:the accumulate reward.

$r = 0$;
$\forall i = 1, 2, ..., K : Q(i) = 0, count(i) = 0$;
for $t = 1, 2, ..., T$ **do**
  randomly choose an arm from 1,2,...,K according to equation (1);
  $v = R(k)$;
  $r = r + v$;
  $Q(k) = \frac{Q(k)*count(k)+v}{count(k)+1}$;
  $count(k) = count(k) + 1$;