

Differential Piracy

Jon Crowcroft
The Computer Laboratory, University of Cambridge
Cambridge, UK
jon.crowcroft@cl.cam.ac.uk

This article is an editorial note submitted to CCR. It has NOT been peer reviewed. The author takes full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT

In all seriousness, Differential *Privacy* is a new technique and set of tools for managing responses to statistical queries over secured data, in such a way that the user cannot reconstruct more precise identification of principles in the dataset beyond a formally well-specified bound. This means that personally sensitive data such as Internet packet traces or social network measurements can be shared between researchers without invading personal privacy, and that assurances can be made with accuracy.

With less seriousness, I would like to talk about Differential *Piracy*, but not without purpose. For sure, while there are legitimate reasons for upstanding citizens to live without fear of eternal surveillance, there is also a segment of society that gets away with things they shouldn't, under a cloak. Perhaps that is the (modest) price we have to pay for a modicum less paranoia in this brave new world.

So, there has been a lot of work recently on Piracy Preserving Queries and Differential Piracy. These two related technologies exploit new ideas in statistical security. Rather than security through obscurity, the idea is to offer privacy through lack of differentiation (no, not inability to perform basic calculus, more the inability to distinguish between large numbers of very similar things).

Categories and Subject Descriptors

C [..]: 2.1 [Packet-switching networks],[Network Communication]
security, privacy

General Terms

General Terms: Algorithms, Performance, Design, Privacy

Keywords

Communications Systems Research. The Internet. Privacy

1. INTRODUCTION

This is about an exciting new development in the world of security called Differential Piracy. Basically, people want to use Aliases to avoid (re-)identification.

So they use random identifiers, rather than truenames, despite what Robert Zimmerman (alias Alias, in Sam Peckinpah's fine elegaic western, Pat Garret and Billy the Kid (case in point)), alias Bob Dylan, and others want you to think.

It's common place, especially amongst internauts, to manage multiple pseudonyms, for example::

Capn' Jack Sparrow Sparrow (in Pirates of the Carribean), Capn' Jim Hawkins, Keith Richard, Keith Richards, and Johnny Depp are all the same person, as we know now.

Dread Pirate Roberts Roberts, in the Princess Bride in fact is a role, taken by different actual people.

Captain Shakespeare Shakespeare (who is definitely in Diagon Ally too) in Neil Gaiman's fine book and film, Stardust has several different roles.

2. BOUNDING MIN AND MAX PIRACY CHARACTERISTICS

We can, of course, at a single leap, bound the Min and Max Piracy Characteristics: Differential Piracy then, lets you get away with it. Databases of pirates are controlled by the *Internaut Coast Guard*, who only allow you to perform certain sorts of queries on the databases -

- Ones which would *not* be permitted, are, for example, ones which would allow you to identify Long John Silver by his extraordinary height, or Capn Hook, by his extraordinary prosthetic limb, etc
- Of course, the counter example is wooden legs and parrots, which would not specifically identify any particular pirate, since all of them have those characteristics. Hence queries like "How many people in Penzance have wooden legs and parrots?" would be perfectly fine.

Of course, Piracy Preserving Queries abound....they are typically of the form "Who ARRRh you?". The reality behind advocacy for *privacy* (to spell out the real message behind this spoof article) is that while there are a small number of occasional legitamate uses of anonimity (e.g. users of ToR in troubled parts of the world where the government are not to be trusted, or whistleblowers), the value of the relative anonimity of today's Internet is that it allows file sharing of copyrighted materials without payment. So privacy begets "piracy". On the other hand, we should be at pains to point out that no-one has actually lost a job or a bean because of Internaut Piracy, whereas real pirates, off the coast of East Africa, actually kill people, and it is insulting to *their* victims, for the media business to use the word

pirate when talking about file uploaders. To misquote Mick Jagger (alias the other half of the Glimmer Twins), “it’s a capital crime, it’s a hanging matter”.

3. ACKNOWLEDGEMENTS

Free Speech, Arrr!

Thanks are due to the Kingston for hosting valuable comments and feedback, and to the CCR Editor for clarifying feedback.

Any feedback will be resoundingly thrashed, and forced to walk the plank.

4. REFERENCES

- Differential Piracy** MSR
<http://research.microsoft.com/pubs/64346/dwork.pdf>
- Pirates of the Carribean** Sparrow
<http://www.imdb.com/title/tt0325980/>
- Stardust** Shakespeare
<http://www.imdb.com/title/tt0486655/>
- The Princess Bride** Roberts
<http://www.imdb.com/title/tt0093779/>
- Peter Pan** Hook
<http://www.gosh.org/gen/peterpan/>
- Pirates of Penzance** Many
<http://www.youtube.com/watch?v=zSGWoXDFM64>