

# Is Content Publishing in BitTorrent Altruistic or Profit-Driven?

Rubén Cuevas (UC3M), Michal Kryczka (Institute IMDEA  
Networks), Angel Cuevas (UC3M), Sebastian Kaune (TU  
Darmstadt), Carmen Guerrero (UC3M) and Reza Rejaie  
(Univ. Oregon)



Rubén Cuevas

[rcuevas@it.uc3m.es](mailto:rcuevas@it.uc3m.es)

University Carlos III of Madrid

# Outline

1. Motivation
2. Data Collection
3. Identifying Major Publishers
4. Incentives of Major Publishers
5. Conclusion



# Outline

1. Motivation
2. Data Collection
3. Identifying Major Publishers
4. Incentives of Major Publishers
5. Conclusion



# Motivation

UK cops arrest six alleged BitTorrent music uploaders  
Surprise new arrests in OiNK investigation

responsible of a

What are the main incentives of content publishers in BitTorrent?

1. Altruism (good citizens) ?
2. Profit-driven incentive ?
3. Any other incentive ?

6					.13 %
7					.58 %
8					.44 %
9					.12 %
10	Facebook	2.28 %	Xbox Live	1.61 %	
Total		78.08 %		82.57 %	

Table 1 - North America - Top Applications by Bytes (Peak Period, Fixed Access)

Source: Sandvine. Fall 2010 Global Internet Phenomena Report



# Outline

1. Motivation
2. Data Collection
3. Identifying Major Publishers
4. Incentives of Major Publishers
5. Conclusion



# Data Collection

- Content Publisher can be identified by
  - **USERNAME** used to register the torrent in a BitTorrent Portal
  - **IP address** used for initially seed the content in the swarm
- We crawl BitTorrent Portals
  - To retrieve the Content Publisher's username (easy)
- We crawl BitTorrent Trackers:
  - To retrieve the Content Publisher's IP address (difficult)
  - To retrieve the IP addresses of the downloaders of a given content



# Our Dataset

Dataset name	BitTorrent Portal	Start Date	End Date	# Torrents (username/ IP addr)	# IP addresses
mn08	Mininova	09-Dec-08	16-Jan-09	- / 20.8K	8.2M
pb09	Pirate Bay	28-Nov-09	18-Dec-09	23.2K / 10.4K	52.9K
pb10	Pirate Bay	06-Apr-10	05-May-10	38.4K / 14.6K	27.3M



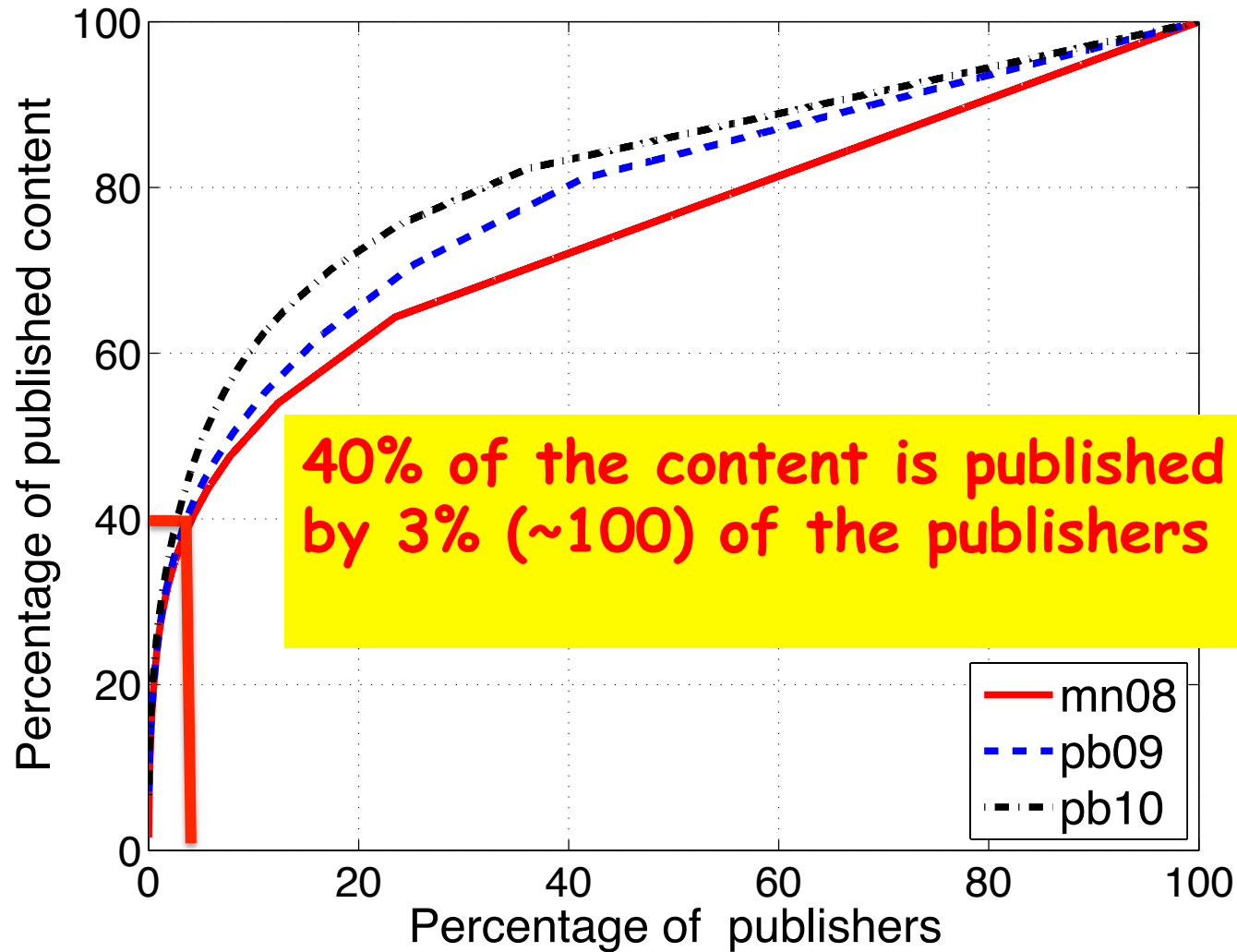
# Outline

1. Motivation
2. Data Collection
3. Identifying Major Publishers
4. Incentives of Major Publishers
5. Conclusion





# Skewness of Contribution



# Identifying major publishers

- Top-100 publishers based on IP addresses
  - 55% use a unique username

"Top" + "Fake" publishers are responsible of  
66% of the content  
75% of the downloads

- 25% use an unique IP
- If we remove the fake usernames -> "Top"
  - 37% content and 50% downloads

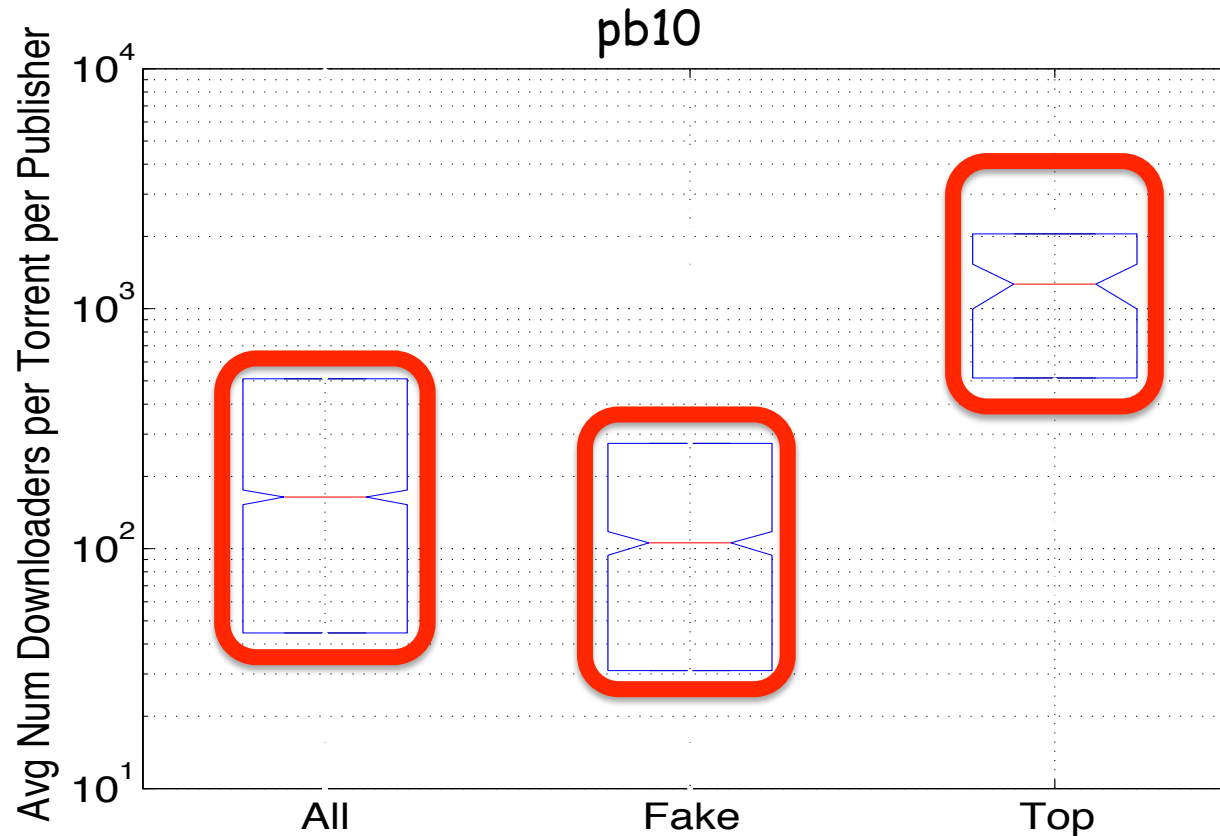


# Signature of Content Publishers

- Do major publishers have any distinguishing features (i.e. signature) ?
- We examine the following characteristics
  - Type of published content
  - **Popularity of published content**
  - Availability and Seeding behavior
    - **Aggregated Session Time per Publisher**



# Popularity of published content



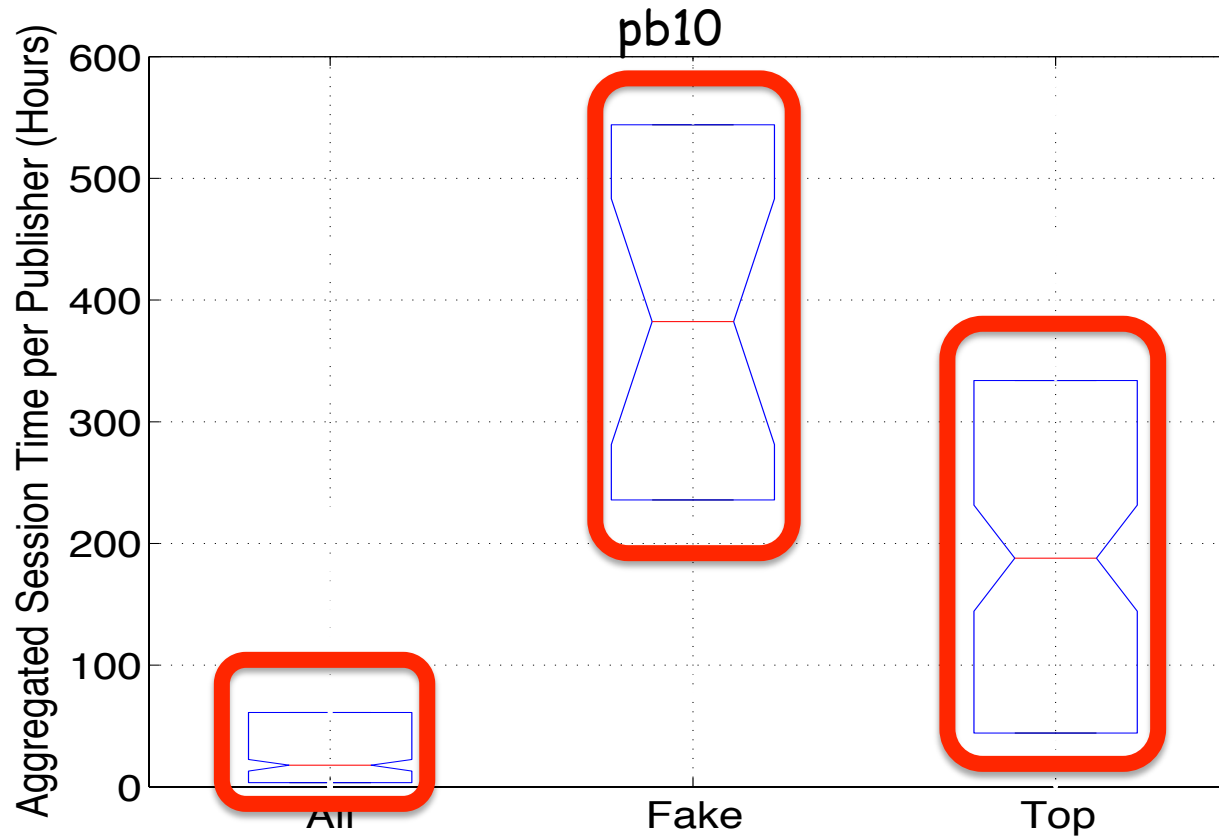
The median popularity of "Top" publishers' torrents is 7 times higher than a typical user ("All")

users quickly report fake content in forums



# Seeding behavior of content publishers

## Aggregated Session Time per Publisher



Not helpful in seeding fake content  
 "Fake" publishers much more available than regular publishers  
 "Top" publishers much more available than regular publishers  
 (more contents seed during longer periods) quires

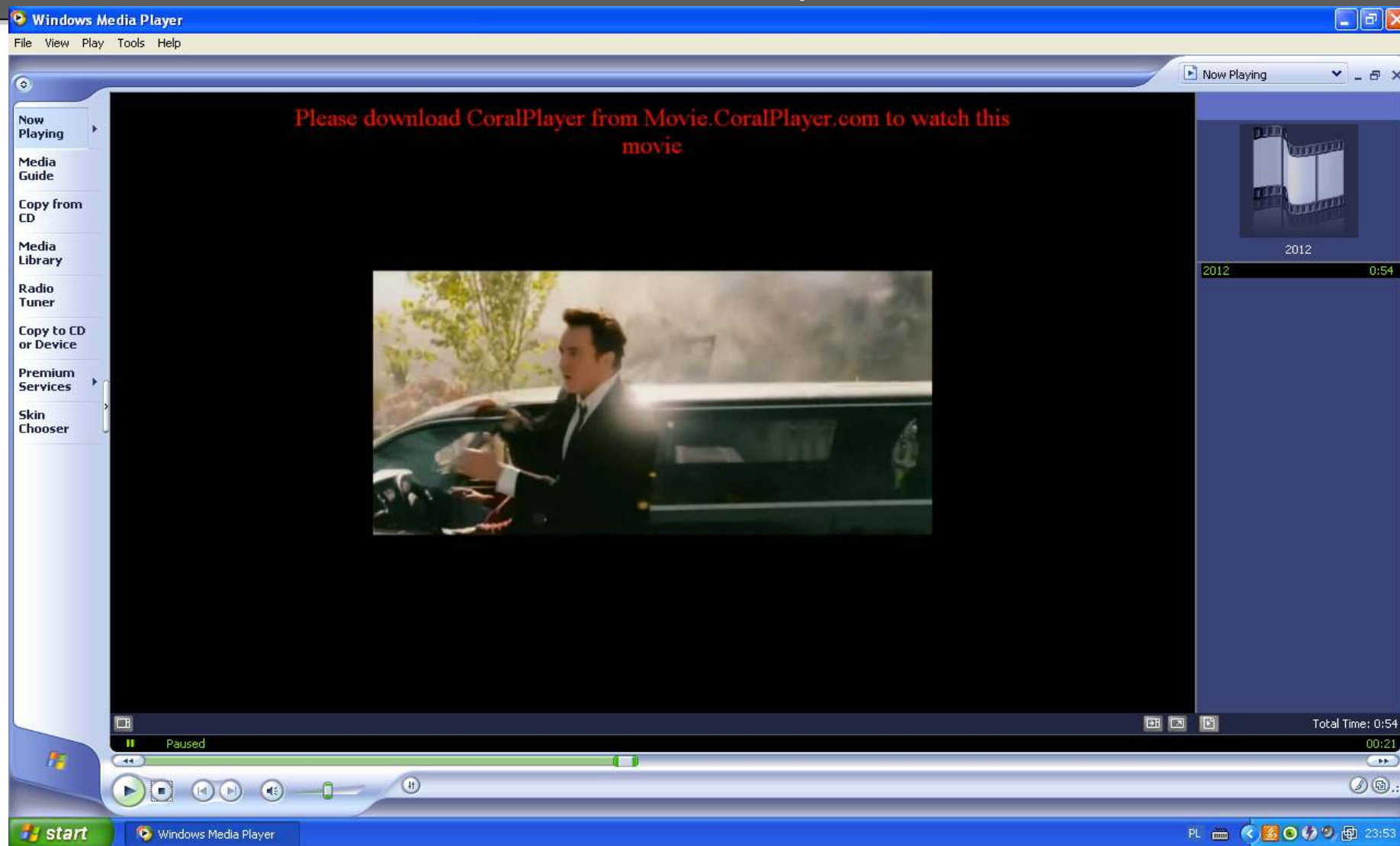


# Outline

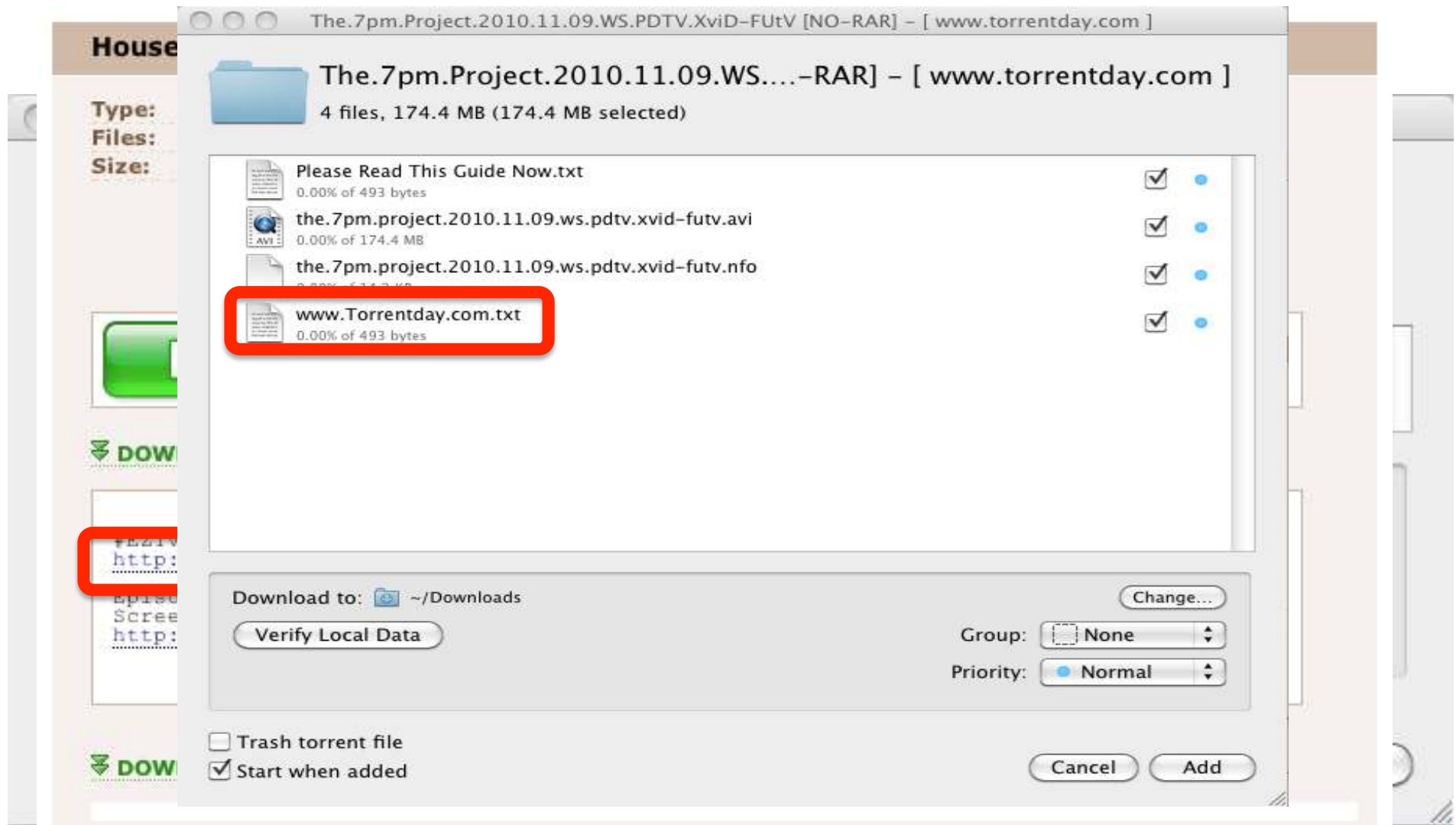
1. Motivation
2. Data Collection
3. Identifying Major Publishers
4. Incentives of Major Publishers
5. Conclusion



# Incentives of "Fake" publishers



# Incentives of "Top" publishers Data Collection





# Incentives of "Top" publishers Profiles

- **Private BT Portals/Trackers (profit-driven)**
  - 18% of the content, 29% of the downloads
  - Financial profit: ads, donations and VIP access
  - Video, audio and software
- **promoted Web Sites (profit-driven)**
  - 8% of the content, 11% of the downloads
  - Mostly hosting (adult) images portals that publish porn content in the Pirate Bay
  - Financial profit: ads
- **Altruistic (major) Publishers**
  - 11.5% of the content, 11.5% of the downloads
  - Light content (music and e-books)
  - Detailed description, asking for help on seeding



# Estimating Publishers' Income of profit-driven publishers

- Promoted Web Sites are fairly profitable and visible

Few of the profit-driven publishers are associated to very valuable web sites (valued up to \$ millions)

	Web Site Value (\$) Min / Median / Max	Web Site Daily Income (\$)	Web Site Daily Visits
Private BT Portals	1K/33K/2.8M	1/55/3.7K	74/21K/1.4M
Promoted Web Sites	24/22K/1.8M	1/51/1.9K	7/22K/772K



# Outline

1. Motivation
2. Data Collection
3. Identifying Major Publishers
4. Incentives of Major Publishers
5. Conclusion



# Conclusion

- 67% of the published content and 75% of

The removal of the financial-driven publishers (e.g. by antipiracy legal actions) may severely affect the popularity of BitTorrent.

If this happens:

<http://bittorrentcontentpublishers.netcom.it.uc3m.es>

*"Will BITTorrent survive as the most popular file-sharing application without these financial-driven publishers?"*

67% of non-pure content and 75% non-pure  
downloads



# Is Content Publishing in BitTorrent Altruistic or Profit-Driven?

Rubén Cuevas (UC3M), Michal Kryczka (IMDEA Networks),  
Angel Cuevas (UC3M), Sebastian Kaune (TU Darmstadt),  
Carmen Guerrero (UC3M) and Reza Rejaie (Univ. Oregon)



## Thanks for your attention!

Rubén Cuevas

[rcuevas@it.uc3m.es](mailto:rcuevas@it.uc3m.es)

University Carlos III of Madrid