

# **Monetary valuation of people's private information**

Marek Kumpošt, Vašek Matyáš

Faculty of Informatics

Masaryk University

**PUMP 2010**

**Dundee, Scotland**

# Introduction of the study

- Usage of online communication tools
- Email or instant messaging used every day
- Network administrators can track users
- Risk of profiling or another analyzes of data
- People can sense the value of such information

# The cover-story & steps

- Sociological study about usability of online communication tools
  - Observed via special software (source code provided, external security audit)
  - Content of messages won't be observed
  - Periodical collection of communication data
  - Financial compensation for participating
  - Placing bids

# Auction principle

- Our budget is limited
- Participants place their bids in an auction
- Selected participants will get the amount required by the first participant „below the line“
- Bidding too high – exclusion from the study
- Bidding too low – discomfort for taking part
- „True value“ of private information

# Organisation of the study

- **First form** (webpage) – do you want to take part?
  - **Academic research**
  - Yes, with a PC/mobile device only (or both)
- **Second form** – partially supporting our cover story
  - Age?, Gender?
  - Own or shared hardware?
  - Level of IT-knowledge?
- **How much money** for being tracked for two weeks
  - email
  - instant messaging
  - all tracking data
- **Second bid – commercial exploitation** (decline, revised bid)
- **Third bid – use by national governments** to improve terrorist activity detection and tracking tools

# Structure of responders

- Intent to participate in the first step (academic research usage of data) of the study – 428 subjects (of 1080 loads)
  - BE(3%), CZ(40.7%), DE(8%),
  - SK(31.8%), EN(17.5%)
- 284 then actually bid (first scenario)
- Those who saw the intro-text and answered
  - will participate – 40 %, (26 % – first scenario)
- Majority of participants were students

# Academic usage (quartiles)

First bids			First bids – males			First bids - females		
email	messaging	all	email	messaging	all	email	messaging	all
10	10	12	10	9.5	12	10	10	15
30	30	50	32.5	25	50	30	35	50
100	100	200	100	100	200	275	150	300

- Quartiles instead of min, max, average values
- The sum of email and messaging is higher than all
- Females (20 % of all participants) bidding higher
- 26 participants (almost 10%) explicitly opt out for the next scenario, but 27% left

# Commercial usage

Academic			Commercial			% increase
email	messaging	all	email	messaging	all	
10	8.3	10.4	10	10	15	22%
20	22.5	40	40	40	50	50%
100	80	150	100	100	200	21%

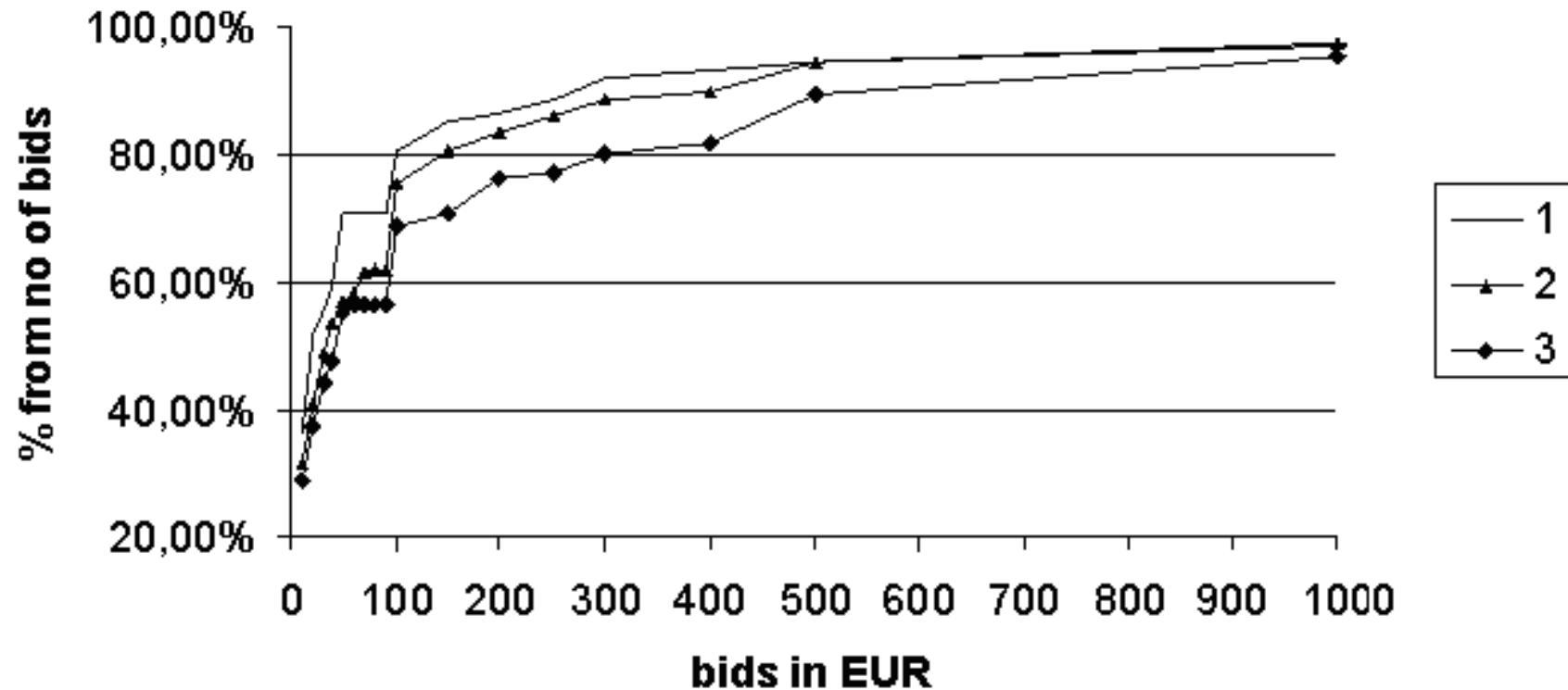
- Medians increased significantly
- 41 participants (18%) explicitly opt out in the next scenario, but 28% actually left



# Usage by governments

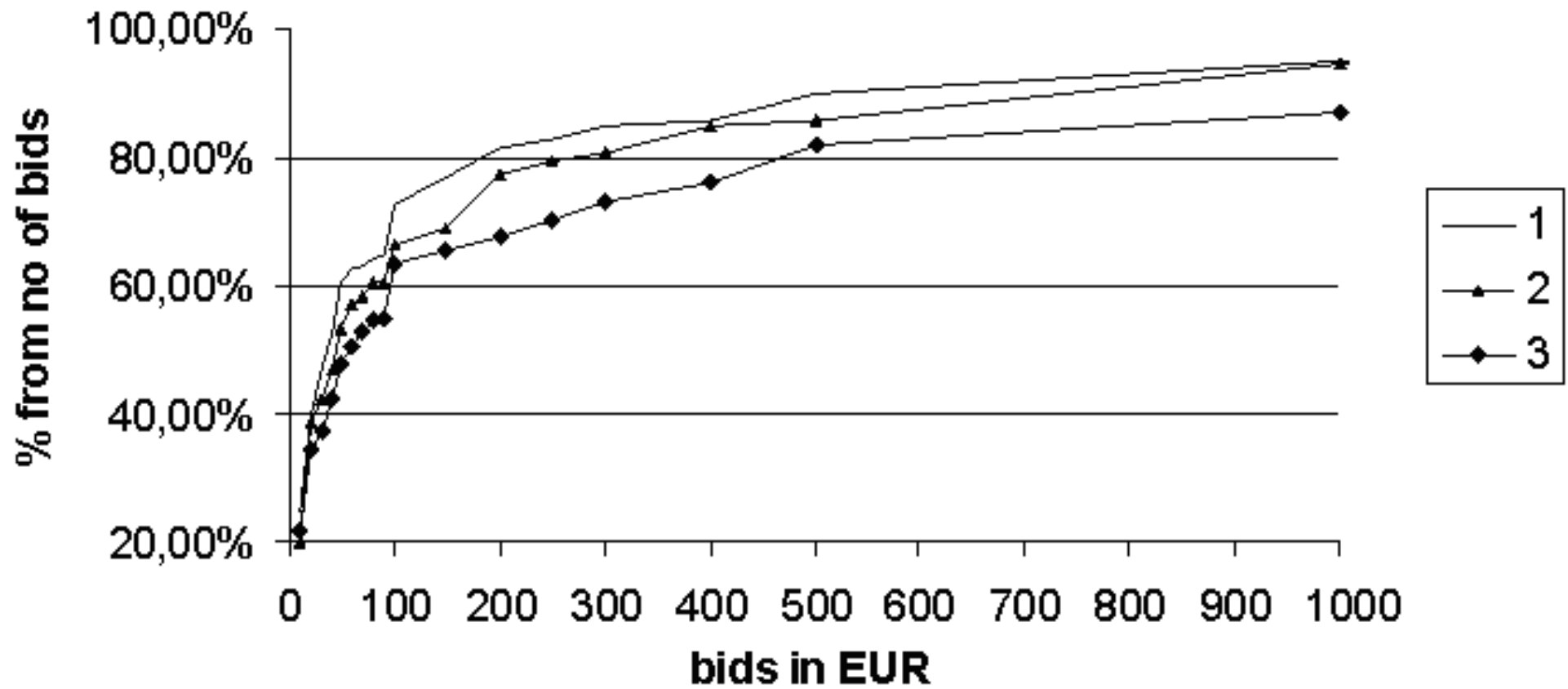
Second bids			Third bids		
email	messaging	all	email	messaging	all
10	10	15	10	10	15
40	40	50	50	50	60
100	100	200	200	200	400

# 1st bid, all scenarios



- Higher differences expected

# 3rd bid, all scenarios



# Distribution of bids

- Are the distributions of bids same or not?
  - and only the level is different...?
- Kolmogorov-Smirnov test
- email vs. online messaging
  - Distribution is the same
  - Average value is 160 vs. 166
- email/online messaging vs. all traffic data
  - Distribution is the same
  - Average value is 160 vs. 288

# Why not participating

- Do not have appropriate HW (3)
- Do not have time (13)
- Do not see the value of such study (12)
- Such study is not ethical (11)
- Do not trust your intentions (6)

# Highlights of the study

- 284 responses for at least the first scenario
  - Responses from more than four countries
- EUR 30 for being tracked (email or instant messaging) for academic purposes
  - EUR 50 for all tracking data
  - No considerable differences between males and females
- Increasing tendency to opt out with changing purpose of tracking
  - 1/10 academic -> commercial usage (real dropout 27%)
  - 1/5 commercial -> governmental usage (real dropout 28%)
- Governmental usage
  - After dropouts, i.e., valuation of all-consenting subjects
  - €50 for one type of data (cf. €40 commercial, €20/25 acad.)
- No significant difference between value of email and other messaging traffic data (statistical tests)
- Females bidding higher in the first scenario, males bidding higher in the second and third scenarios

Thanks for your attention