



# 582519 Scientific Writing for MSc in Computer Science: Ethics of writing

Lecture 4, 24.9.2013

Tiina Niklander



# Last week's task: How did it go?

---

-- fill comments here --

Paper layout  
List of references and its structure  
Table of contents  
First pages



# Honesty and ethics of science

---

- Science is based on trust
  - Researchers are honest
  - Research is ethical
- Honesty
  - Ideas, results, and conclusions presented are new
  - Assumptions are not presented as facts
  - Truth is not misrepresented
- Violation of these norms of behaviour is a serious matter!



# Reviewing / Referring

---

- Criticising and analysing others' work
  - Research article
  - Theses, reports, seminar work
- Important part of the scientific process
- Should be objective and fair
- Referee statement / reviewer statement
  - Analysis of the contents of the article
  - Not evaluation of the author or his/her working place



# Reviewing / Referring (2)

---

- Conflicts of interest in referring
  - Personal or employment relationship
  - Competition for an appointment
  - Conflicting ideas/opinions
  - Too similar research interests/ideas
  
- A suspicion of non-objectiveness  
=> refuse the referee task



# Reviewing / Referring (3)

---

- Confidentiality
  - As referee / reviewer you are not allowed to publish the articles or parts of them
  - Not even allowed to show to others, if not an assisting referee
- Totally out of line
  - Suggest rejection, and then
  - Use the ideas/results as a basis for referee's own work



# Plagiarism

---

- The most obvious form of unethical behaviour in science
- Attempt to get honour of work that somebody else has done
- Using existing material without citing to the original work, or citing it inadequately
- Intentional and unintentional plagiarism
  - disregard of the norms of ethical behaviour
  - ignorance of those norms



# Plagiarised material

---

- Plagiarised material can be
  - ideas
  - results
  - text, pictures, figures, tables
  - whole articles or part of them
- Sources of the plagiarised material
  - published articles
  - web pages
  - newsgroup articles
  - e-mail message, ....





# Forms of plagiarism

---

- Use of others' ideas or results without acknowledgements
- Direct copying of material
  - without citing the source
  - citing the source, but not indicating the exact quotation
- Using material with inappropriate or inadequate citations
  - What part of the text the citation covers?



## Forms of plagiarism (2)

---

- Copying the structure of a source
- Using pictures
  - Copying of a unique picture (by scanning or re-drawing it) always plagiarism
  - If modified from its original form (even translated), the source must be indicated
- Self-plagiarism



# How not to plagiarize (source: University of Toronto)

<http://www.writing.utoronto.ca/advice>

---

- The purpose of any paper is to show your own thinking, not to create a patchwork of borrowed ideas.
- Citations are used to demonstrate that you know what is going on in your field of study.
- In academic papers, you need to keep mentioning authors and pages and dates to show how your ideas are related to those of the experts.
- As you read the text summarize useful points in your own words. (Do not copy to draft – that causes trouble!)

[www.writing.utoronto.ca/advice/using-sources/how-not-to-plagiarize](http://www.writing.utoronto.ca/advice/using-sources/how-not-to-plagiarize)



# Self-plagiarism

---

- Copying/reusing one's own text (previous work)
  - Exception: an extended/complete version of a conference/workshop paper that is published in a journal
- Multiple articles based on same results
  - Publishing improper unless full cross-referenced
  - Simultaneous submissions to different publication forums must be disclosed



# Self-plagiarism (2)

---

- Against good ethics
  - Even the description of the background of the work should always be rewritten
  - A member of a research group cannot use texts of the other group members
- Question of copyright
  - Publication forum has usually the rights for the presentation
  - Author has the rights for his/her ideas!



# Authorship

---

- Each author of an article should have some contribution to the contents of it
  - Implementing an algorithm not enough
  - Giving feedback not enough

=> mentioned in Acknowledgements
- Preferred: contribution to ideas, experimentation and analysis
- Authors must give their permission for authorship
- Being a member of a research group does not automatically give the authorship



# Authorship (2)

---

- Postgraduate studies
  - Student and supervisor together
  - Supervisor should not publish alone results that the student has obtained
  - Student should not publish the results without consulting the supervisor
- Order of the authors
  - Alphabetical order by last names
  - The author with the biggest contribution first



# Publishing on web pages

---

- Articles published only in authors' or research groups' web pages
  - Intention to publish as a conference, workshop or journal article
  - Can be considered by a publication forum as published
  - Usually not refereed
- Articles already published in a conference, workshop or journal
  - Not always allowed
  - When allowed, must typically include the copyright notice of the publisher





# Detecting plagiarism

---

- Usually it is easy to tell when someone has copied text
- Language changes
  - Language style varies
  - Fluency of text varies
  - Terminology changes
    - especially if copied from different places and sources
  - Words not a normal part of the writer's style and vocabulary



## Detecting plagiarism (2)

---

- Digital copies of articles make copying easy
  - **BUT** they can be easily found also by the referees and teachers!
- Our university and department **do not tolerate plagiarism**, i.e. copying the work of others
- Technical tools are used for detecting plagiarism and checking the originality of submitted reports, seminar works and theses!



# Example of detection tools: Turnitin

Originality GradeMark PeerMark

mobile DTN  
BY HASAN ISLAM

turnitin 46% INDEX OUT OF 0

Match Overview

Rank	Source	Percentage
1	ftp.cs.duke.edu Internet source	7%
2	people.ee.ethz.ch Internet source	5%
3	www.rcf.usc.edu Internet source	5%
4	ieeexplore.ieee.org Internet source	4%
5	web.cs.wpi.edu Internet source	2%
6	www.sigcomm.org Internet source	2%
7	Cauligi S. Raghavendra... Publication	2%
8	eprints.biblio.unitn.it Internet source	1%
9	epubl.ltu.se Internet source	1%
10	lib.tkk.fi Internet source	1%
11	ceng.usc.edu Internet source	1%
12	M. Mauve. "A survey on..." Publication	1%
13	Yinglie Cao. "A positi..." Publication	1%

**Abstract** Intermittently connected mobile networks are a collection of wireless mobile nodes. In such network there is no assurance that a complete path exist between source and destination. Traditional routing protocol is unable to deliver message between nodes as the nodes are highly mobile and the path between nodes is unstable in nature. These networks can be generalized as Delay Tolerant Network. Many real networks like wildlife tracking sensor network, interplanetary networks, and nomadic communities network are fall into such category of networks. However, though the connectivity in such network is unreliable, researchers have proposed a lot of routing algorithms in the domain of Delay Tolerant Network. In this article, I present several existing routing strategies in intermittently connected networks and finally focus on Spray and Wait. Routing algorithms which outperform all existing schemes in terms of average message delivery delay and number of transmissions per message. It is shown that Spray and Wait is highly scalable in sparse networks and its performance is close to optimal algorithm.

## I. INTRODUCTION

Intermittently connected mobile networks are a collection of wireless mobile nodes where there is no guarantee that a completely connected path exist at any given time. In such networks, a path from source to destination is unstable, and unreliable and may change frequently due to the nodes are highly mobile in nature. Moreover, a path may break after it has been discovered or even while being discovered. Intermittently connected networks are fall into the category of Delay Tolerant Network (DTNs) that are characterized by the lack of connectivity and instantaneous end to end path.

In areas where there is little or no communication infrastructure or the existing infrastructure is expensive, wireless mobile users may able to communicate through the formation of an ad hoc network as such network can be built on the fly without any preexistent infrastructure. In such network, each node acts as a client as well as server. In ad hoc network, each source node forwards packets to destination through other intermediate node though the source and destination

hoc network routing algorithms, such as AODV [3], DSDV [4], DSR [5] would not work.

In hostile and rapidly changing mobile ad hoc networks, the establishment of connectivity between handheld devices or between vehicles is a challenging task as there exist no preexisting networking system and the topology of such network change frequently without any prior notice. In [6], two different approaches in routing are distinguished: topology-based and position-based routing. The former routing protocols use the information of the link which exists in the network topology to transfer packet from source to destination. They can be further divided into three categories: proactive, reactive, and hybrid routing scheme.

Proactive algorithms utilize basic routing algorithms of ad hoc networks such as distance vector routing (DSDV [4]) or link state routing (OSLR [7]). They maintain available path information about the network whether the path is currently in use or not. In Reactive routing protocols (e.g., DSR, AODV), only the currently used routing information is maintained. In Hybrid ad hoc routing protocols (ZRP [8]), the local proactive routing and global reactive routing are combined for better performance and scalability. However, all the mentioned routing protocols consist some inevitable drawbacks. Proactive algorithms occupy a significant amount of bandwidth if the network topology changes sporadically. Reactive algorithms incurred delays to deliver the first packet as they are to maintain the routing information while the routes are in use. A survey and comparison of topology based routing schemes are presented in [9], [10].

Position based routing schemes extinguish some of the limitations of topology based routing by introducing additional information about the position of nodes that are participating in communication. Each node determines its own position through the use of GPS [11]. Some other methods are also found in [12]. A location service is used by the sender to determine the specific position of the destination. Before participating in communication, each mobile node registers its current position with the location



# Cut and paste plagiarism

promote their products and services to customers, in monitoring inventory of products and in shopping in stores through RFID. As per industrial studies mobile marketing can be accepted by customers if the marketing benefits them.

Researchers suggested that some improvements in MCDs such as functionality of the keyboards and screens of MCDs, and the prolonged existence of the devices, can increase the number of features on the devices for an increased number of consumers in our society.

## A. PRIVACY AND SECURITY CONCERNS

“Location privacy” is the ability to control the disclosure of one’s location information to other parties.

According to researchers benefits of location-based services are only one side of the coin, the issues of customer privacy is the other side of it. As mobile telephony becomes very common which enables location-based services to spread outside closed environments, which raises the serious issues of customer privacy in relation to the building of location based technologies and services. Customer privacy concerns are about control of personal and private information on the mobile devices and about fear of frequent incidents on the devices of likely identity theft and intrusion on the privacy of consumers.

Researchers further discuss about role of privacy activists, that they have cited many fundamental issues such as the mismanagement and marketing of information on citizens and

of America only governs the federal government and the financial and health care industries in information and in rights to privacy but there are no explicit privacy protection rights available in the Constitution of the United States, in other American industries generally. Therefore consumers are dependent on privacy policies of other industries.

Whereas according to European Directive 95/46/EC, for privacy protection information has to be processed fairly and lawfully, collected for explicit and legitimate purposes and not further processed in a manner inconsistent with such purposes, not excessive in relation to the collected or processed purposes, current, and in a form that permits identity of consumers no longer than necessary.

Even in the presence of such a coherent, well implemented and very protective legislation than in America, European consumers still have to be dependent on privacy and security practices in other industries like in America.

Therefore according to researchers, poorly defined and executed legislation policies for controlling the use of location based information, are aggravating privacy issues in America. Federal legislation defined in the Telecommunications Act of 1996, location-based information about a mobile consumer as customer proprietary network information (CPNI) for completing calls for customers but not for marketing products and services to them. Further it was not clearly defined in this Act for the carrier or the provider was the form of opt-in or opt-out by customers for the products and services. However, because of this inconsistency and vagueness in legislation the



# What about this one? Plagiarism or not?

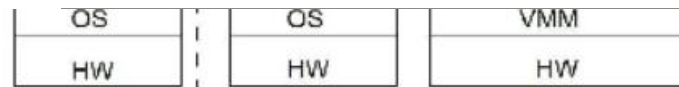


Figure 4.1 CloneCloud system model [4]

Another limitation of CloneCloud is its inability to migrate native state and to export unique native resources remotely. That's why it focuses on migration at execution points where no native state in the stack or the heap needs to be collected and migrated. CloneCloud neither virtualizes access to native resources that are not virtualized already or are not available on the cloud.

Despite of CloneCloud's limitations, Chun et. al. noticed significant speed-ups when they tested their prototype with three different applications. Their prototype delivered up to 21.2x speed-up without programmer involvement.

## B. AlfredO

Giurgiu et. al. [5] have also researched how to dynamically distribute applications between the cloud and mobile devices. Their approach has many similarities and differences with CloneCloud. They use AlfredO to distribute an application between a mobile phone and a server. It is based on OSGi, which has been used to decompose and loosely couple Java applications to software modules.

Giurgiu's et. al. approach does not yet do resource profiling automatically so it has to be done manually. They compose a

architecture of cloud storage system. Cloud storage system is based on cooperation between multiple devices, many application domains, and many service forms. There exist many different cloud storage service platforms, but they are usually complex and incompatible. Therefore, Zeng. et. al. want to propose a layered and generalized architecture of cloud storage. It consists of five layers: network and storage infrastructure, storage management, metadata management, storage overlay, and service interface (Figure 5.1).

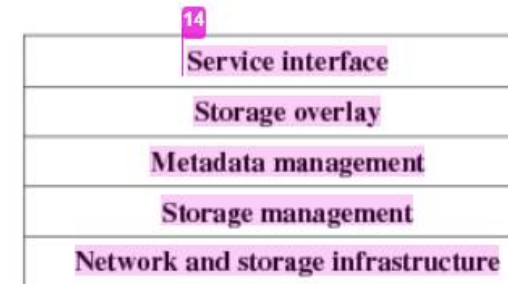


Figure 5.1: Cloud storage layered model [6]

According to Zeng et. al., cloud storage system is actually an implementation of storage as a service which is expected to be available, reliable, cooperative, scalable, secure, concurrent and economical. When constructing cloud storage system requirement analysis, capacity prediction and performance planning deployment, verification, distribution,



# Plagiarism detection at University of Helsinki

---

- Using URKUND
  - submissions via email or moodle
- Used in this course for all submissions (even the drafts)
  - group leaders will tell your the email address
- More information see:  
<http://blogs.helsinki.fi/alakopsaa/for-student/?lang=en>
- National guidelines  
[http://www.tenk.fi/sites/tenk.fi/files/HTK\\_ohje\\_2012.pdf](http://www.tenk.fi/sites/tenk.fi/files/HTK_ohje_2012.pdf)



(IT PROTECTS YOUR COPYRIGHT!)



I SAVE MY DOCUMENT INTO THE DATABASE

It's advisable to save the document into the Urkund database: this way you protect your own copyright!

A student



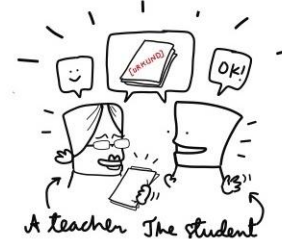
And then:



A report is sent:



Teacher's screen:



A teacher The student



UNIVERSITY OF HELSINKI

Read more at: ↘

<http://blogs.helsinki.fi/alakopsaa>





# Consequences of plagiarism in general

---

- Article is not published
- Lost of reputation
- Lost of a job
- Even jail
- ...





## Consequences of plagiarism (2) for students

---

- Thesis/report/seminar work is not accepted
- Course is failed
- Course must be retaken
- ...
- Personal meeting with Head of the Department, Dean of the Faculty and/or Rector of the University
- Student will lose the study right, i.e., is thrown out of the university



# Avoiding plagiarism

---

- Take model of others work (outlining, writing, citing), but don't copy them
- Cite your sources
- Make it clear what is the origin of your text
  - Distinguish what you state and what the others have stated
  - Paraphrase the ideas of others



## Avoiding plagiarism (2)

---

- Show direct quotations with quotation marks (even direct translations) and proper citations
- Make your own pictures, tables, etc.
- Give in the list of references only those articles that you have read and cited



# Paraphrasing

---

- Writing the ideas of some other authors in your own words
  - Without changing the original meaning
  - Paragraph structure not the same
  - Sentence structure not the same
  - Words not too similar to the original
- Requires that you understand what the others have stated
- Paraphrasing can also help to understand a difficult text



# Paraphrasing (2)

---

- Basic process
  - Read the source and understand it
  - Write/paraphrase the contents with your own words
  - Check your version against the original
  - Revise if necessary
- Different approaches of paraphrasing (writing)
  - **Read the source, put it a side, and do not look at it, while writing**
  - Read the source and take notes, have a break, use the notes while writing
  - Paraphrase while looking at the source



# Paraphrasing (3)

---

- Key steps of the paraphrasing process
  - Changing the structure
    - Paragraph structure
    - Sentence structure
    - Keep just the main and most relevant ideas
  - Changing the words
    - Keep specialised words and terminology, i.e., shared language
    - Find alternative words and expressions for other words and phrases
      - Thesauruses, dictionaries



# Paraphrasing (4)

---

- Paraphrasing is an iterative process
  - Start by changing the structure, not the words
  - Change the words
  - Make further changes to the structure
  - ...
- Several iterations may be needed!



# Misrepresentation

---

- Occurs when
  - the results obtained improperly described and/or overstated
  - the value of previous work diminished
- Description of the results
  - Accurate, precise, correct, truthful
  - All restrictions explained
  - Both negative and positive results reported
  - Detailed enough so that experiments can be repeated





# Misrepresentation (2)

---

- Unintentional mistake is not misrepresentation
- The most serious form: totally incorrect statement(s)
- Other forms:
  - Underestimating previous work
  - Emphasising own results (even preliminary or somehow restricted)
  - Omitting unsuccessful experiments and their results
  - Omitting the change history of web documents



# Sources

---

- Books
  - Zobel, J., Writing for Computer Science: The art of effective communication. Springer-Verlag, Singapore, 1997.
  - Lester, J.D., Writing Research Papers: A complete guide. 7<sup>th</sup> edition. HarperCollins College Publishers, New York, 1993.
  - Barrass, R., Scientists Must Write: A guide to better writing for scientists, engineers and students. Chapman & Hall, London, 1995.



## Sources (2)

---

- Online guides, for example,
  - Writer's Handbook of University of Wisconsin-Madison (USA)
  - Online writing lab of University of Purdue (USA)
  - Writing skills guide of the Royal Melbourne Institute of Technology (Australia)
  - Justin Zobel's page with links to Technical writing and research ethics



**Submission address for me:**

**Tiina.Niklander.hy@analyysi.urkund.fi**



# Time management

---

- How to balance studies, sleeping and hobbies/family life / free time?
  - any ideas?
- Timetable for studies
- Study plan: which courses and when?

## Decisions, skills, tools

Cristallising goals

→ motivation, plans, followup

Selection and prioritisation of tasks

- based on values
- systematically
- clear subgoals in terms of Contents, quality, schedules
- persistence, courage

Long term resourcing  
vs instant pleasure

Control of own and shared  
Ways of work, e.g. Reporting,  
Feedback, splitting work

Removal of disturbance  
- materials, tools, peaceful work  
Environment, notes available for  
No delay access

Self confidence

Distributing one's own resources  
according to own values

Feeling of acceptance,  
value rapport

Process skills, leadership,  
subordinance skills, expert role

**self**

Management  
Of your own  
work

Feeling of control  
Satisfaction of work  
Quality of work  
Keeping deadlines

Cyclic development of  
life management

## Threats, hinders

Stress, anxiety, exhaustion

Feeling of insufficiency

Avoiding larger tasks

Procastination

Unclearly stated goals

Decisionmaking difficulty ←  
Partitioning of work, commitment

Unrealistic expectations →  
Difficulty estimating quality of work  
Failure to resource

Difficulty reporting  
→ failure to receive advisory

Difficulty completing task  
← expectations on advisor's push