HOW TO AVOID PLAGIARISM?

Tiina Niklander
AMICT 2011, Petrozavodsk

Table of content

- Plagiarism definition
- Copy detection tools
- Turnitin
  - Tool overview
  - Criticism
- Avoiding unintentional plagiarism
- Conclusion

Plagiarism definition

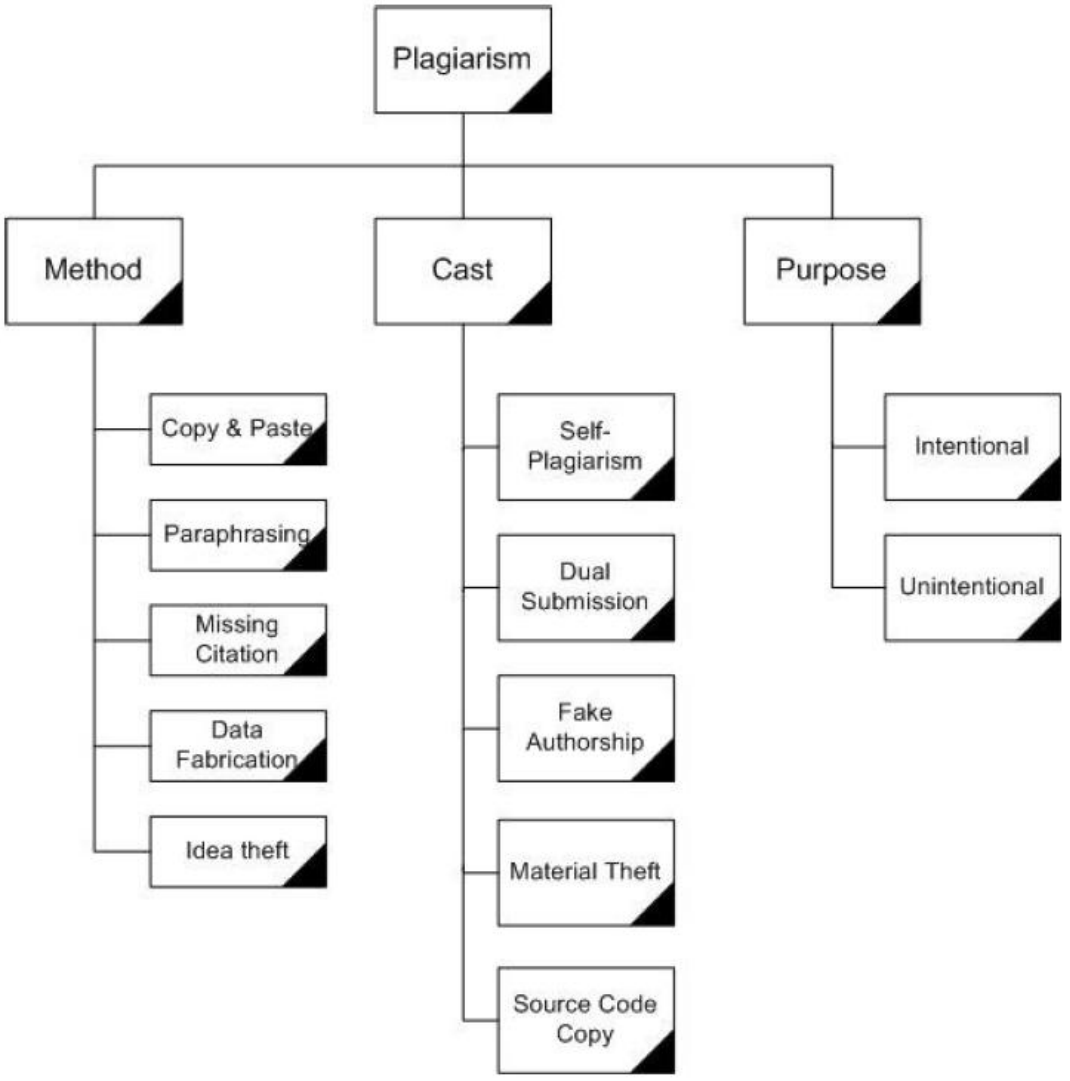
- ”Plagiarism is the use of another’s thoughts, or work, without acknowledgement or permission. In plagiarism, one author takes another’s idea and presents it as his/her own.” (Smith & Wren 2010)
- “Plagiarism originated from the Latin word *plagiarius* meaning the theft of words as well as slaves” (Gu & Brooks, 2008)
- “The word plagiarism comes from Latin word *plagium* which means kidnapping” (Imran, 2010)

N. Smith & K. Wren: Ethical and Legal Aspects Part 2: Plagiarism—“What Is It and How Do I Avoid It?”. *Journal of PeriAnesthesia Nursing, Vol 25, No 5 (October), 2010: pp 327-330*

Q. Gu & J. Brooks: Beyond the accusation of plagiarism. *System 36 (2008) 337–352,* Elsevier.

Types of plagiarism in research

- Method
  - how it is implemented
- Cast
  - the form of plagiarism
- Purpose
  - the motivation behind



N. Imran: Electronic Media, Creativity and Plagiarism. *SIGCAS Computers and Society, Volume 40, No. 4, December 2010*

Copy detection tools

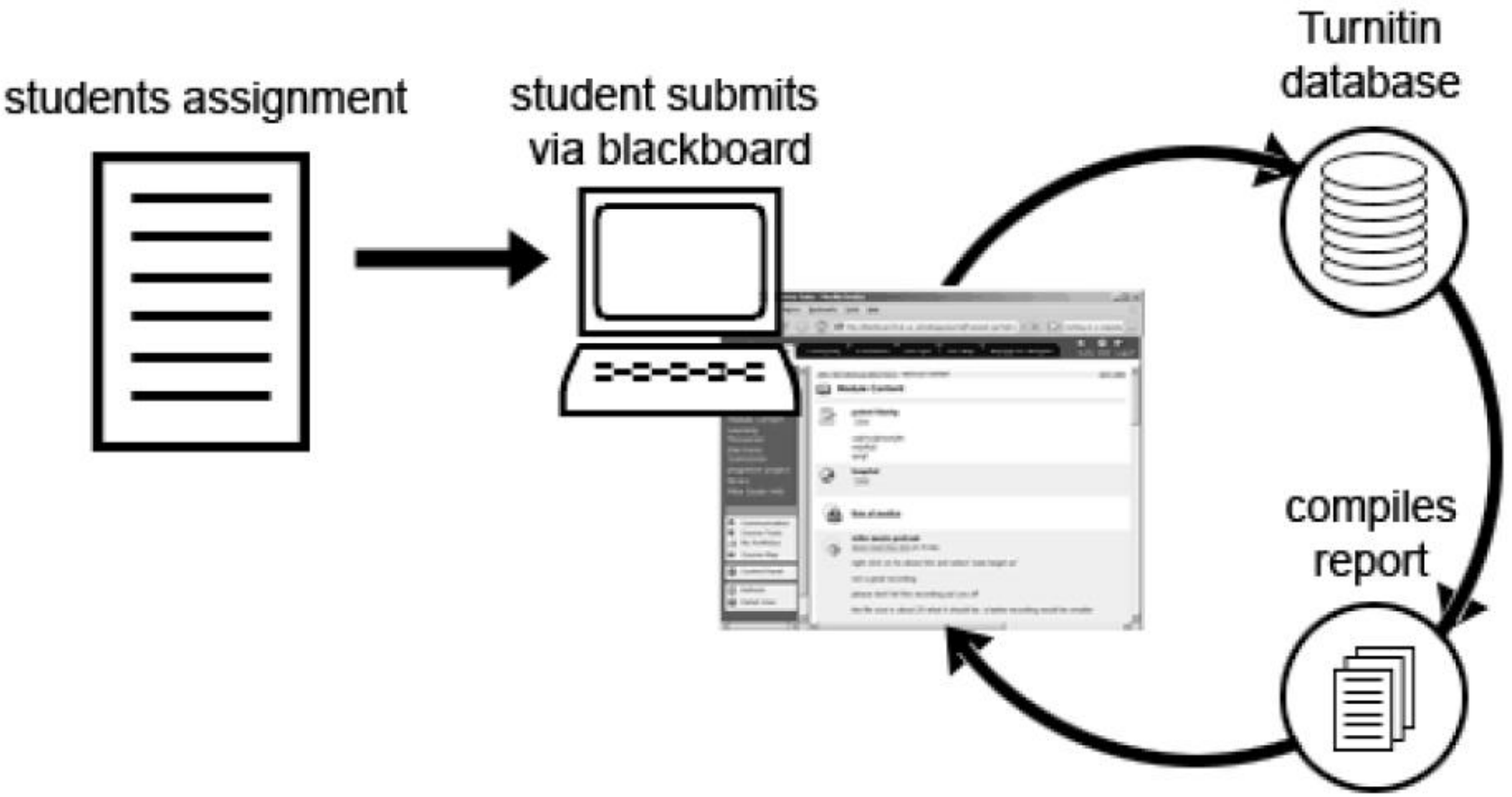
- Automated tools to detect similarities in compared documents
- Statistical methods – most common
  - N-gram, Lancaster word pairs, word frequencies, compression metrics – fingerprints
  - Euclidean distance, cosine function, Jaccard measure, Dice measure
- Several tools available
  - De facto standard : turnitin.com
  - Others: Eve2, CopyCatchGold, WordCheck, Glatt, Moss, JPlag

R. Lukashenko, V. Graudina & J. Grundspenkis: Computer-Based Plagiarism Detection Methods and Tools: An Overview. *International Conference on Computer Systems and Technologies - CompSysTech'07, ACM, 2007,* IIIA.18-1 - IIIA.18-6

Turnitin

- Web-based tool
- Own database of articles and web pages
- No ACM & IEEE? - at least claimed in 2008 by Kaner and Fiedler
- ”Ignoring the commonly used words, the software looks for matches of strings of eight to ten words” (Jones, 2008)

K. Jones:
Practical Issues for Academics Using the Turnitin Plagiarism Detection Software.
International Conference on Computer Systems and Technologies - CompSysTech’08, ACM, 2008, IV.1-1 - IV.1-5

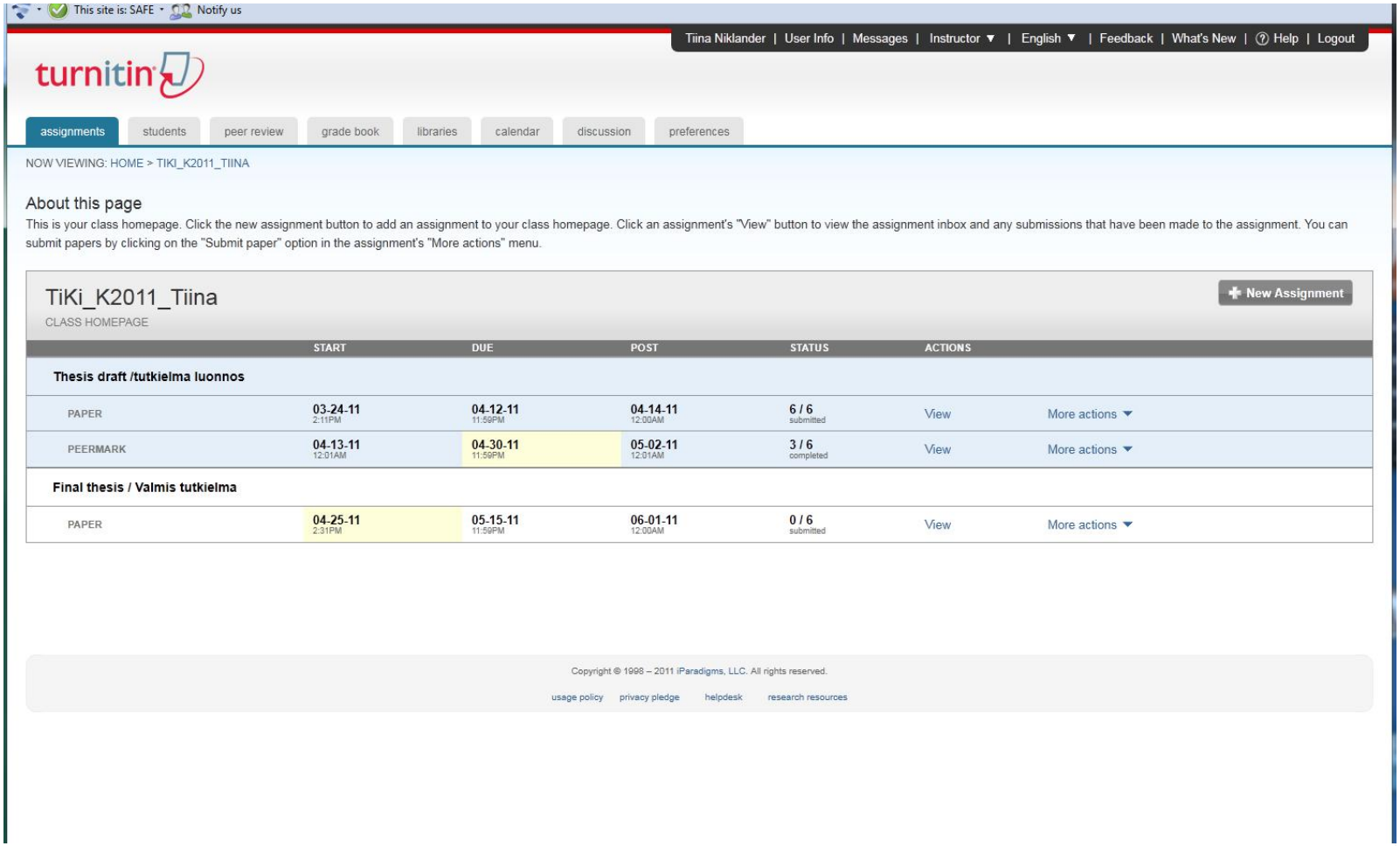


Tool overview

Student submits her work, possibility for teacher submission

Can be integrated to course platforms like blackboard or moodle

Fig 1 from K. Jones: Practical Issues for Academics Using the Turnitin Plagiarism Detection Software. *Intern. Conference on Computer Systems and Technologies - CompSysTech'08, ACM, 2008,* IV.1-1 - IV.1-5

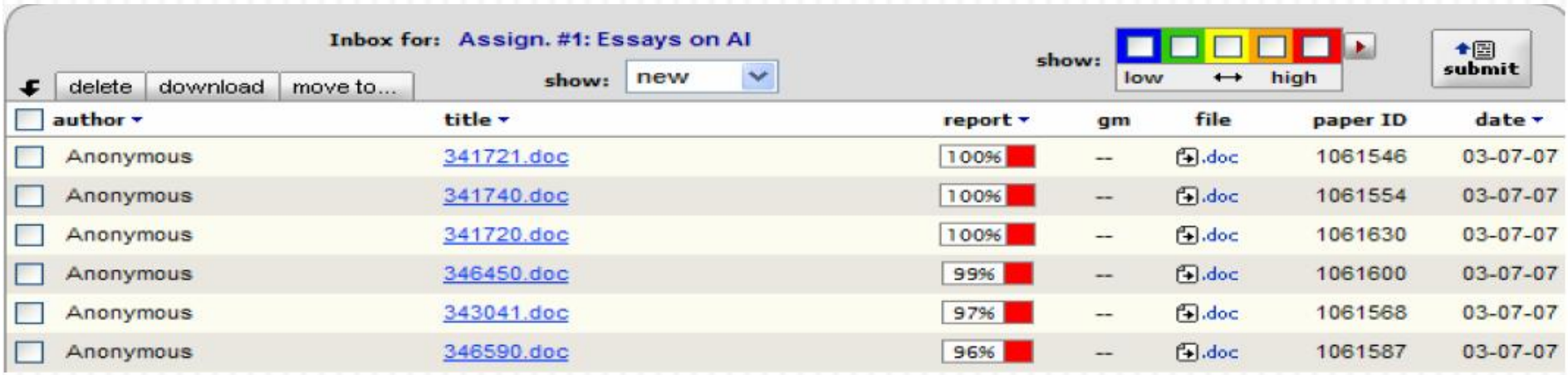


Turnitin.com

Example course with two separate paper submission tasks and peer review of the first one

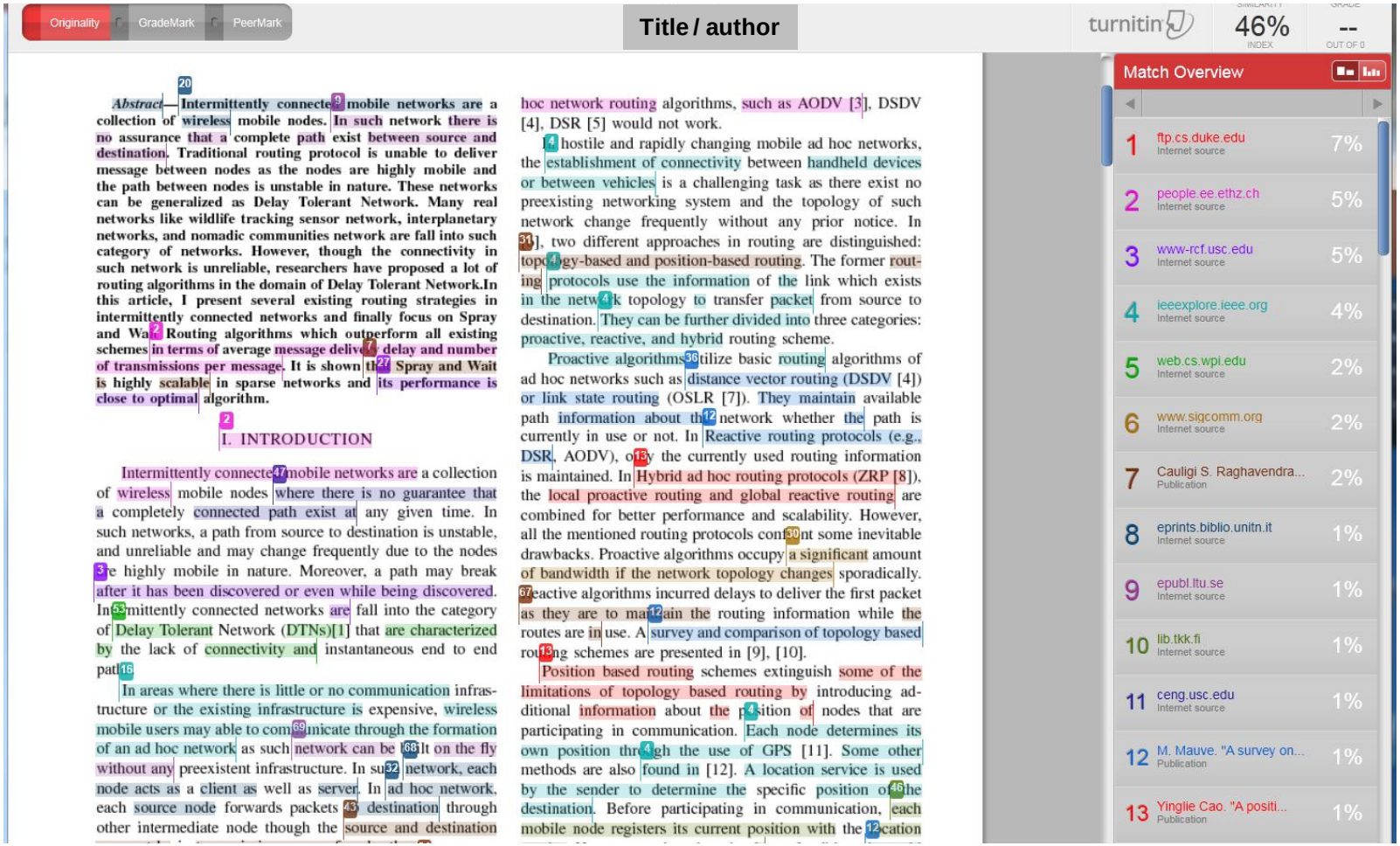
Originality measure

- Traffic lights (blue, green, yellow, orange, red) + percentage
- How much of the submission matches with existing material



author	title	report	gm	file	paper ID	date
Anonymous	341721.doc	100%	--	.doc	1061546	03-07-07
Anonymous	341740.doc	100%	--	.doc	1061554	03-07-07
Anonymous	341720.doc	100%	--	.doc	1061630	03-07-07
Anonymous	346450.doc	99%	--	.doc	1061600	03-07-07
Anonymous	343041.doc	97%	--	.doc	1061568	03-07-07
Anonymous	346590.doc	96%	--	.doc	1061587	03-07-07

Fig 2 from K. Jones: Practical Issues for Academics Using the Turnitin Plagiarism Detection Software. *International Conference on Computer Systems and Technologies - CompSysTech'08, ACM, 2008,* IV.1-1 - IV.1-5



Turnitin originality report window

Submitted text in the big picture, total originality value on top right, matched articles on the right with similarity percentage

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

Cut and paste plagiarism

Clearly fullfils the definitions of plagiarism:

direct copying and no citations.

(colors show different sources)

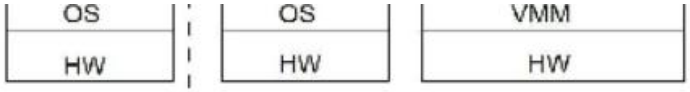


Figure 4.1 CloneCloud system model [4]

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).

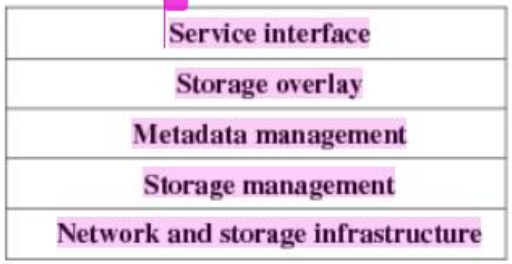


Figure 5.1: Cloud storage layered model [6]

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,

What about this one?

There are citations in place and the names of the authors of the articles are mentioned.

Notice – one main colour per (sub)section !

Criticism about Turnitin

- Byrocratic decisions based on the originality percentage should not be done automatically using a boundary value for similarity
  - Asian students ’caught’ too easily
  - English-speaking students can rephrase better, but still copy idea
  - Student copying from single paper might remain under the limit
- Teaching is more important than punishment, because normally not intentional by the students

Q. Gu & J. Brooks: Beyond the accusation of plagiarism. *System 36 (2008) 337–352,* Elsevier.

Avoiding unintentional plagiarism

- "Avoiding plagiarism does not need to be difficult or require an in-depth knowledge of copyright law." (Smith&Wren 2010)
- Use proper referencing
- Paraphrase properly
- Summarize in own voice (with citation)
- *Understand the source information* (Imran 2010)

N. Smith & K. Wren: Ethical and Legal Aspects Part 2: Plagiarism—"What Is It and How Do I Avoid It?". *Journal of PeriAnesthesia Nursing, Vol 25, No 5 (October), 2010: pp 327-330*

N. Imran: Electronic Media, Creativity and Plagiarism. *SIGCAS Computers and Society, Volume 40, No. 4, December 2010*

Explanations given by students

R. Comas-Forgas & J. Sureda-Negre: Academic Plagiarism: Explanatory Factors from Students' Perspective. *J Acad Ethics (2010) 8:217–232 . Springer.*

- Aspects and behaviour of students
  - bad time management,
  - personal shortcomings when preparing assignments,
  - the elevated number of assignments to be handed in, etc.
- Opportunities conferred by ICT to locate, copy and paste information
- Aspects related to professors-lecturers and/or the characteristics of the subject-course
  - lecturers who show no interest in their work,
  - eminently theoretical subjects and assignments, etc.

Conclusion

- Common sense
- Do not steal ideas
- Give credit to the right persons and papers
- Tools help teachers in checking

- ”The very nature of research is to examine and build upon previous findings and to confirm, deny, or expand into new areas.” (Smith & Wren 2010)

References

- R. Comas-Forgas & J. Sureda-Negre: Academic Plagiarism: Explanatory Factors from Students' Perspective. *J Acad Ethics (2010) 8:217–232 . Springer.*
- Q. Gu & J. Brooks: Beyond the accusation of plagiarism. *System 36 (2008) 337–352,* Elsevier.
- N. Imran: Electronic Media, Creativity and Plagiarism. *SIGCAS Computers and Society, Volume 40, No. 4, December 2010*
- K. Jones: Practical Issues for Academics Using the Turnitin Plagiarism Detection Software. *International Conference on Computer Systems and Technologies - CompSysTech'08, ACM, 2008,* IV.1-1 - IV.1-5
- R. Lukashenko, V. Graudina, J. Grundspenkis: Computer-Based Plagiarism Detection Methods and Tools: An Overview. *International Conference on Computer Systems and Technologies - CompSysTech'07, ACM, 2007,* IIIA.18-1 - IIIA.18-6
- N. Smith & K. Wren: Ethical and Legal Aspects Part 2: Plagiarism—"What Is It and How Do I Avoid It?". *Journal of PeriAnesthesia Nursing, Vol 25, No 5 (October), 2010: pp 327-330*