Research Methods in computer science

Spring 2025

Lecture 10

Omprakash Gnawali

February 19, 2025

Agenda

CS Experiments

HW5

Types of Papers (mechanical)

- Technical Reports
  - Project description
  - Research paper
- Conference
- Journal
- Magazine

Find out what type your group and community writes.

Which papers are more important?

Conference
Journal
Magazine

Impact factor
CORE ranking
DBLP

What makes a paper more important than others?

Types of Papers (purpose)

Research Paper

Survey Paper

Tutorial

Technical Report

- E.g., NIST, Other Orgs

White Paper

Vision Paper

Challenge Paper

Publications – Looking Ahead

Blogs?

Facebook?

Twitter?

LinkedIn?

GitHub?

YouTube?

arXiv?

…..

Citation and References

Clean! Clean! Clean!

(esp. for websites, links, datasheets)

Consistency! Consistency! Consistency!

Examples

ALIZAI, M. H., WIRTZ, H., KIRCHEN, B., VAEGS, T., GNAWALI, O., AND WEHRLE, K. 2011. TinyWifi: Making Network Protocol Evaluation Portable Across Multiple Phy-Link Layers . In *WiNTECH '11: Proceedings of the Sixth ACM International Workshop on Wireless Network Testbeds, Experimental evaluation and Characterization*.

BHATTI, S., CARLSON, J., DAI, H., DENG, J., ROSE, J., SHETH, A., SHUCKER, B., GRUENWALD, C., TORGERSON, A., AND HAN, R. 2005. MANTIS OS: An Embedded Multithreaded Operating System for Wireless Micro Sensor Platforms. *Mobile Networks and Applications 10*, 4 (Aug.), 563–579.

BROUWERS, N., LANGENDOEN, K., AND CORKE, P. 2009. Darjeeling, a feature-rich VM for the Resource Poor. In *SenSys '09: Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM, New York, NY, USA, 169–182.

BURRI, N., VON RICKENBACH, P., AND WATTENHOFER, R. 2007. Dozer: ultra-low power data gathering in sensor networks. In *IPSN '07: Proceedings of the 6th international conference on Information processing in sensor networks (2007-05-25)*. ACM, 450–459.

CHALLEN, G. W., WATERMAN, J., AND WELSH, M. 2010. IDEA: Integrated Distributed Energy Awareness for Wireless Sensor Networks. In *MobiSys '10: Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, New York, NY, USA, 35–48.

CHEN, Y., GNAWALI, O., KAZANDJIEVA, M., LEVIS, P., AND REGEHR, J. 2009. Surviving Sensor Network Software Faults. In *SOSP '09: Proceedings of 22nd ACM Symposium on Operating Systems Principles*.

CHIPARA, O., LU, C., BAILEY, T. C., AND ROMAN, G.-C. 2010. Reliable Clinical Monitoring Using Wireless Sensor Networks: Experiences in a Step-down Hospital Unit. In *SenSys '10: Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. ACM, New York, NY, USA, 155–168.

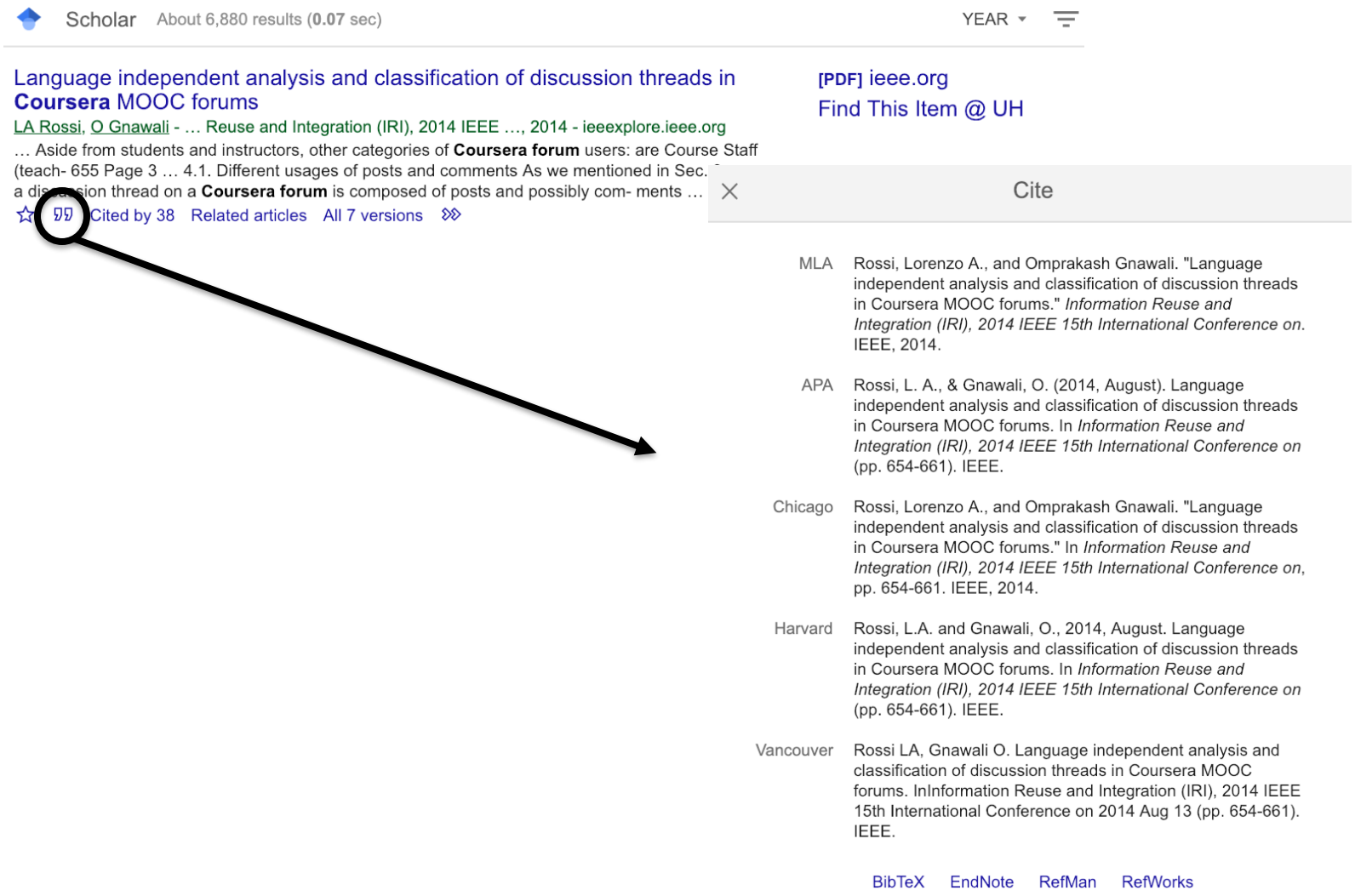
CONTIKICOLLECT. 2010. Contiki Collect Memo. `http://comments.gmane.org/gmane.os.contiki.devel/5790`.

References

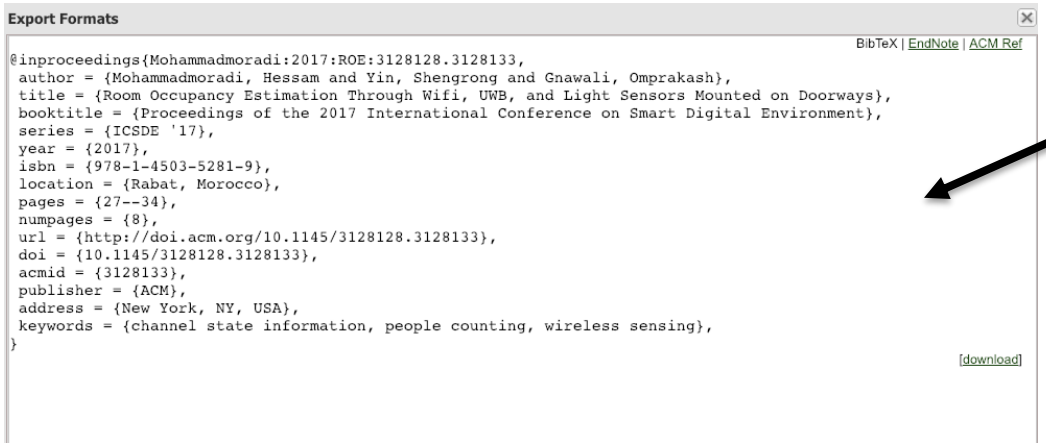
Can take a long time to format references.

Is it worth it?

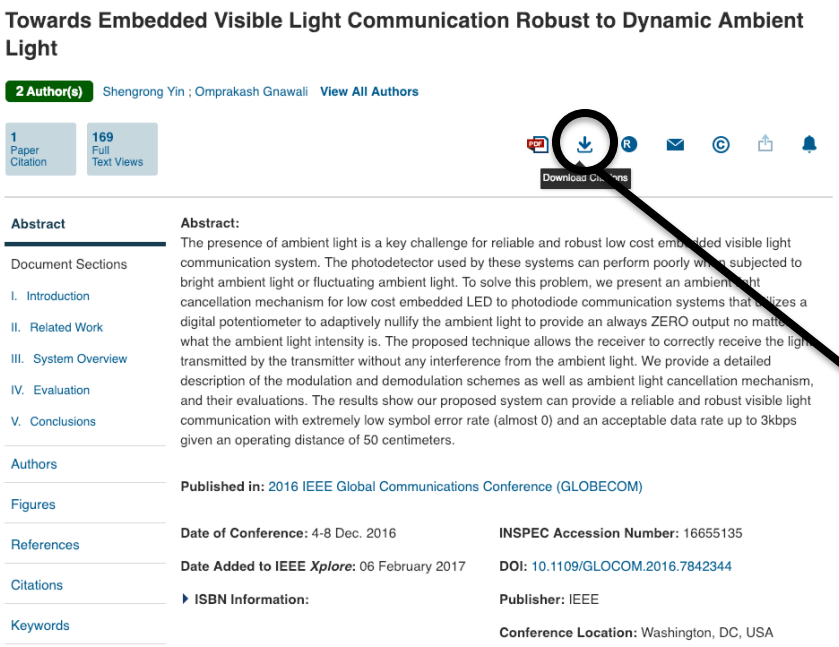
Citations – Google Scholar



Citations – ACM DL



Citations - IEEE



Towards Embedded Visible Light Communication Robust to Dynamic Ambient Light

2 Author(s) Shengrong Yin ; Omprakash Gnawali **View All Authors**

1 Paper Citation

169 Full Text Views

Download Citations

Abstract

Document Sections

I. Introduction

II. Related Work

III. System Overview

IV. Evaluation

V. Conclusions

Authors

Figures

References

Citations

Keywords

Abstract:
The presence of ambient light is a key challenge for reliable and robust low cost embedded visible light communication system. The photodetector used by these systems can perform poorly when subjected to bright ambient light or fluctuating ambient light. To solve this problem, we present an ambient light cancellation mechanism for low cost embedded LED to photodiode communication systems that utilizes a digital potentiometer to adaptively nullify the ambient light to provide an always ZERO output no matter what the ambient light intensity is. The proposed technique allows the receiver to correctly receive the light transmitted by the transmitter without any interference from the ambient light. We provide a detailed description of the modulation and demodulation schemes as well as ambient light cancellation mechanism, and their evaluations. The results show our proposed system can provide a reliable and robust visible light communication with extremely low symbol error rate (almost 0) and an acceptable data rate up to 3kbps given an operating distance of 50 centimeters.

Published in: 2016 IEEE Global Communications Conference (GLOBECOM)

Date of Conference: 4-8 Dec. 2016

Date Added to IEEE *Xplore*: 06 February 2017

ISBN Information:

INSPEC Accession Number: 16655135

DOI: 10.1109/GLOCOM.2016.7842344

Publisher: IEEE

Conference Location: Washington, DC, USA

```
@INPROCEEDINGS{7842344,
author={S. Yin and O. Gnawali},
booktitle={2016 IEEE Global Communications Conference (GLOBECOM)},
title={Towards Embedded Visible Light Communication Robust to Dynamic Ambient Light},
year={2016},
volume={},
number={},
pages={1-6},
keywords={demodulation;free-space optical communication;interference suppression;light emitting diodes;optical modulation;photodetectors;photodiodes;dynamic ambient light fluctuation;robust low-cost embedded visible light communication system reliability;photodetector;photodiode communication system;low-cost embedded LED;digital potentiometer;modulation scheme;demodulation scheme;ambient light cancellation mechanism;distance 50 cm;Receivers;Photodiodes;Robustness;Modulation;Prototypes},
doi={10.1109/GLOCOM.2016.7842344},
ISSN={},
month={Dec},}
```

References - 1

[36] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *Proc. IEEE Conf. CVPR*, Jun. 2013, pp. 3025–3032.

[37] Z. Cui, W. Li, D. Xu, S. Shan, and X. Chen, "Fusing robust face region descriptors via multiple metric learning for face recognition in the wild," in *Proc. IEEE CVPR*, Jun. 2013, pp. 3554–3561.

[38] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, "An empirical evaluation of deep architectures on problems with many factors of variation," in *Proc. 24th ICML*, 2007, pp. 473–480.

[39] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 2032–2047, Nov. 2009.

[40] A. Krizhevsky, "Learning multiple layers of features from tiny images," 2009.

[41] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *J. Mach. Learn. Res.*, vol. 9, pp. 1871–1874, Jul. 2008.

[42] D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. IEEE Conf. CVPR*, Jun. 2012, pp. 3642–3649.

[43] K. Sohn, G. Zhou, C. Lee, and H. Lee, "Learning and selecting features jointly with point-wise gated Boltzmann machines," in *Proc. 30th ICML*, 2013, pp. 217–225.

[44] K. Yu, Y. Lin, and J. Lafferty, "Learning image representations from the pixel level via hierarchical sparse coding," in *Proc. IEEE Conf. CVPR*, Jun. 2011, pp. 1713–1720.

[45] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

References - 2

[1] N. H. Motlagh, M. Bagaa, and T. Taleb, "UAV-based IOT platform: A crowd surveillance use case," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 128–134, 2017.

[2] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1123–1152, 2016.

[3] (2017, April) Homeland security in united states. [Online]. Available: https://en.wikipedia.org/wiki/Homeland_security

[4] C. C. Haddal and J. Gertler, "Homeland security: Unmanned aerial vehicles and border surveillance." DTIC Document, 2010.

[5] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[6] R. K. Sharma and D. B. Rawat, "Advances on security threats and countermeasures for cognitive radio networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1023–1043, 2015.

[7] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1550–1573, 2014.

[8] J. Xu, K. Li, L. Duan, and R. Zhang, "Proactive eavesdropping via jamming over HARQ-based communications," in *IEEE Global Communications Conference (GLOBECOM)*, 2017.

[9] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

Citations

- Systems and artifacts
  - Generally immediately after the names

    AnguLoc [5] is better than SideLoc[6].

- Narrative
  - Generally at the end of the sentence

  Researchers have made a lot of progress in this field in the last five years [6]

This solution is scalable as the number of transmitting anchors can be small and can be scheduled in different time slots. With the usage of inter-anchor concurrency, solutions like AnguLoc [1] managed to make it more efficient. However, this architecture is not cost-effective as described earlier.

In addition to that, new light-emitting technologies, such as LEDs, become more popular and accessible, enabling new perspectives for optical wireless communication [7], [8]. Finally, the increasing interest and exploration of the

Bad form:

[1] presents a new technique.

Citation Format

Number: [n]

Author / year: [Gnawali et al. 2020]

In text without [] or (), e.g., Gnawali et al. proposed a new technique.

Consult the instruction for your conference or journal. Number [n] format common in our fields.

Hypothesis

“A hypothesis (plural hypotheses) is a proposed explanation for a phenomenon. For a hypothesis to be a scientific hypothesis, the scientific method requires that one can test it. Scientists generally base scientific hypotheses on previous observations that cannot satisfactorily be explained with the available scientific theories. Even though the words "hypothesis" and "theory" are often used synonymously, a scientific hypothesis is not the same as a scientific theory. A working hypothesis is a provisionally accepted hypothesis proposed for further research, in a process beginning with an educated guess or thought.”

-- wikipedia

Hypothesis in Engineering

The hypothesis-model is good for research where you want to understand how something works, but I think it is ill-suited for capturing the full scope of engineering research. After all, in engineering, you're primary goal is not to learn something about how the world works, but rather to **change how the world works**! So, instead of a hypothesis on how something works, I'd put up existing gaps in the ability to do something as a working basis. That will then put a focus on your research result as an extension of technical capabilities. In order to evaluate your research results, you would then have to show that your results actually close the existing gap.

Of course, also engineering research needs to understand something about how existing things work in order to be able to create something new. Hypotheses **are suitable** in engineering to clarify these preliminary things. In your case, you state that *"the existing 'role-based access control' of MS-Windows does not solve some problems"* - that sounds like a perfect hypothesis to test for. But verifying this hypothesis is certainly not the key step in your research, and maybe it has already been done previously. That's why I'd recommend not to focus on a hypothesis as the basis for engineering research (though one might use them to clarify preliminaries), but focus on **identified gaps in current technical abilities**.

-- silvado, Aug 26, 2013 on stackexchange

https://academia.stackexchange.com/questions/12156/hypothesis-for-an-engineering-oriented-research-thesis

Hypothesis in Engineering

Effectively, what you are doing is *development* of existing research, rather than designing something *de novo*. The notion of a research hypothesis is therefore somewhat inappropriate to such work, and you wouldn't write a paper describing this work specifying a definitive "hypothesis."

Instead, you'd write the paper emphasizing that your model does something "better," "faster," "more securely," or specifying whatever other accomplishments advance your work from the previous state of affairs. Your thesis should then show how that is accomplished, and give some evidence thereof.

-- aeismail, Aug 25, 2013

https://academia.stackexchange.com/questions/12156/hypothesis-for-an-engineering-oriented-research-thesis

Hypothesis and Engineering Thesis

Because engineers invent rather than discover, does an engineering thesis need a hypothesis?... because invention is a more tightly directed activity than discovery; and the two are not mutually exclusive anyway…uppose your project involves using Artificial Neural Networks (ANNs), in conjunction with appropriate hardware, to sort good apples from bad. The hypothesis for this project may be, 'It is possible to sort good apples from bad using ANNs and suitable hardware'…. Suppose that on completing your project, you discovered that the system you had devised works well with green apples, but not with red ones. You would have discovered new knowledge and would be able to suggest a revised hypothesis as the starting point for further investigation. Your own project would have demonstrated the correctness of a hypothesis like 'It is possible to sort good green apples from bad green apples, with an accuracy of better than 90%, using ANNs and suitable hardware'.

http://thesishub.org/does-an-engineering-thesis-need-a-hypothesis/

“We build new XYZ” – not sufficient.

We can call it a hypothesis or not. We need to know what questions we are trying to answer.

Sample Hypothesis

Only an extraordinarily skilled attacker can break into our firewall. [?]

The firewall accepts all well-formed packets and sessions, and handles malformed packets and sessions as documented in the firewall's manual.

From Sean Peisert and Matt Bishop

Most of the time our questions are related to what improves some system and the nature of those improvements.

We need to make measurements.

Metric

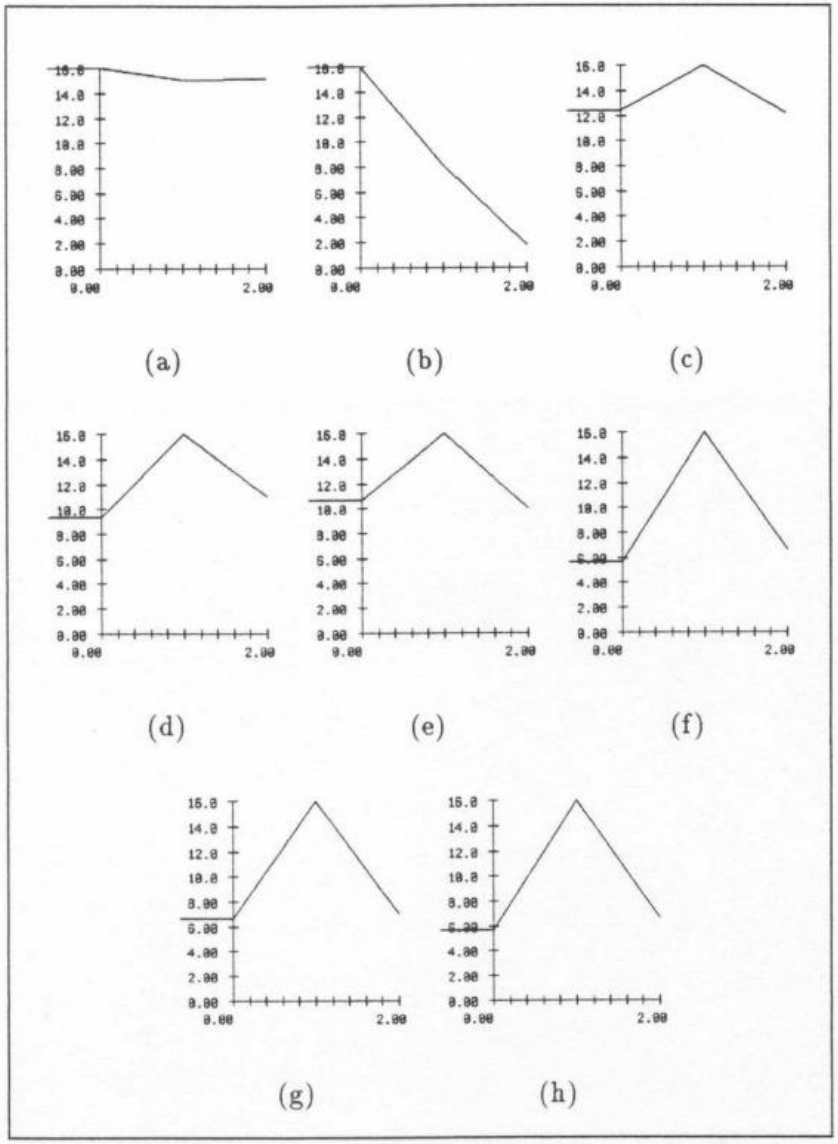
Why do we want to measure?

What to measure?

Eigenfaces for Recognition

[Turk ’91]

“We have developed a near-real-time computer system that can locate and track a subject’s head, and then recognize the person by comparing the characteristics of the face to those of known individuals.”



Scenarios and metrics from [Turk '91]

Figure 9. Results of experiments measuring recognition performance using eigenfaces. Each graph shows averaged performance as the lighting conditions, head size, and head orientation vary—the *y*-axis depicts number of correct classifications (out of 16). The peak (16/16 correct) in each graph results from recognizing the particular training set perfectly. The other two graph points reveal the decline in performance as the following parameters are varied: **(a)** lighting, **(b)** head size (scale), **(c)** orientation, **(d)** orientation and lighting, **(e)** orientation and size (#1), **(f)** orientation and size (#2), **(g)** size and lighting, **(h)** size and lighting (#2).

The Anatomy of a Large-Scale Hypertextual Web Search Engine

[Brin and Page ’98]

What hypothesis, scenarios, and metrics should we expect to see in this paper?

5 Results and Performance

The most important measure of a search engine is the quality of its search results. While a complete user evaluation is beyond the scope of this paper, our own experience with Google has shown it to produce better results than the major commercial search engines for most searches. As an example which illustrates the use of PageRank, anchor text, and proximity, Figure 4 shows Google's results for a search on "bill clinton". These results demonstrates some of Google's features. The results are clustered by server. This helps considerably when sifting through result sets. A number of results are from the whitehouse.gov domain which is what one may reasonably expect from such a search. Currently, most major commercial search engines do not return any results from whitehouse.gov, much less the right ones. Notice that there is no title for the first result. This is because it was not crawled. Instead, Google relied on anchor text to determine this was a good answer to the query. Similarly, the fifth result is an email address which, of course, is not crawlable. It is also a result of anchor text.

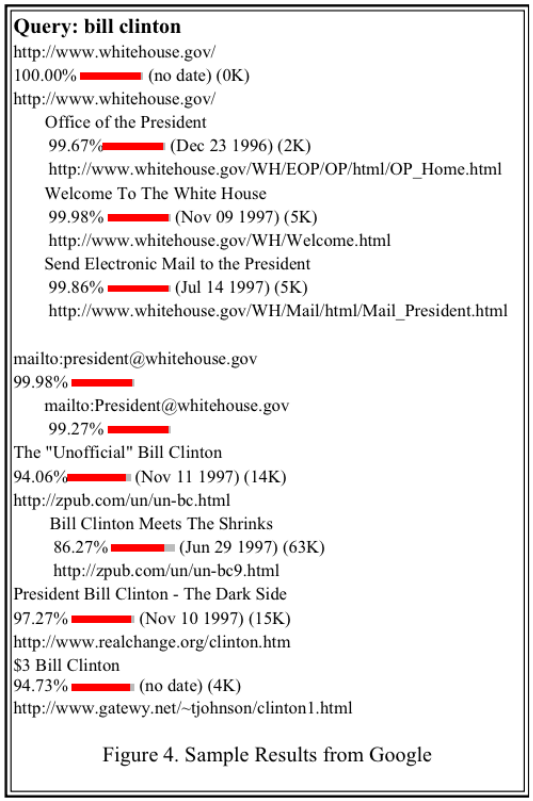


Figure 4. Sample Results from Google

All of the results are reasonably high quality pages and, at last check, none were broken links. This is largely because they all have high PageRank. The PageRanks are the percentages in red along with bar graphs. Finally, there are no results about a Bill other than Clinton or about a Clinton other than Bill. This is because we place heavy importance on the proximity of word occurrences. Of course a true test of the quality of a search engine would involve an extensive user study or results analysis which we do not have room for here. Instead, we invite the reader to try Google for themselves at http://google.stanford.edu.

[Brin and Page '98]

Storage Statistics	
Total Size of Fetched Pages	147.8 GB
Compressed Repository	53.5 GB
Short Inverted Index	4.1 GB
Full Inverted Index	37.2 GB
Lexicon	293 MB
Temporary Anchor Data (not in total)	6.6 GB
Document Index Incl. Variable Width Data	9.7 GB
Links Database	3.9 GB
Total Without Repository	**55.2 GB**
Total With Repository	**108.7 GB**

Web Page Statistics	
Number of Web Pages Fetched	24 million
Number of Urls Seen	76.5 million
Number of Email Addresses	1.7 million
Number of 404's	1.6 million

Table 1. Statistics

[Brin and Page '98]

Why did the authors decide to report these measurements?

Metrics/Experiments?

Accurately Initializing Real Time Clocks to Provide Synchronized Time in Sensor Networks

CTP: An Efficient, Robust, and Reliable
Collection Tree Protocol for Wireless Sensor Networks

On the Effectiveness of Energy Metering on Every Node

Surviving Sensor Network Software Faults

Metrics from Classification Research

Classification Accuracy
Logarithmic Loss
Area Under ROC Curve
Confusion Matrix
Classification Report
Precision
Recall
F1-Score

Partly from https://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/

Metrics from Regression Research

Mean Absolute Error

Mean Squared Error

R^2

Partly from https://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/

Metrics from Systems Research

Reliability

Latency

Coverage

Energy

HW5 – Metrics

List of metrics from the related papers.

Define the metrics.

Observations about common and uncommon metrics.