

# Academic Writing for Computer Science

Master Studies, 2021/2022

Dr Veljko Pejović  
Veljko.Pejovic@fri.uni-lj.si



University of Ljubljana  
Faculty of Computer and  
Information Science

Based on materials by Janet Kayfetz,  
Columbia University/UCSB  
and observations by MOMENT lab, UCSB

# Writing

“If you want to be a writer, you must do two things above all others: read a lot and write a lot. There’s no way around these two things that I’m aware of, no shortcut.”

Stephen King from **On Writing** (2000)



# Academic Writing

- **Main goal:** disseminating knowledge
- Constructing a logical argument around your research
- Position your work within a larger body of knowledge
- Promote your research among the scientific community

This lecture is not  
about English

This lecture is not  
about writing  
popular literature



# Academic Writing



WWW.PHDCOMICS.COM



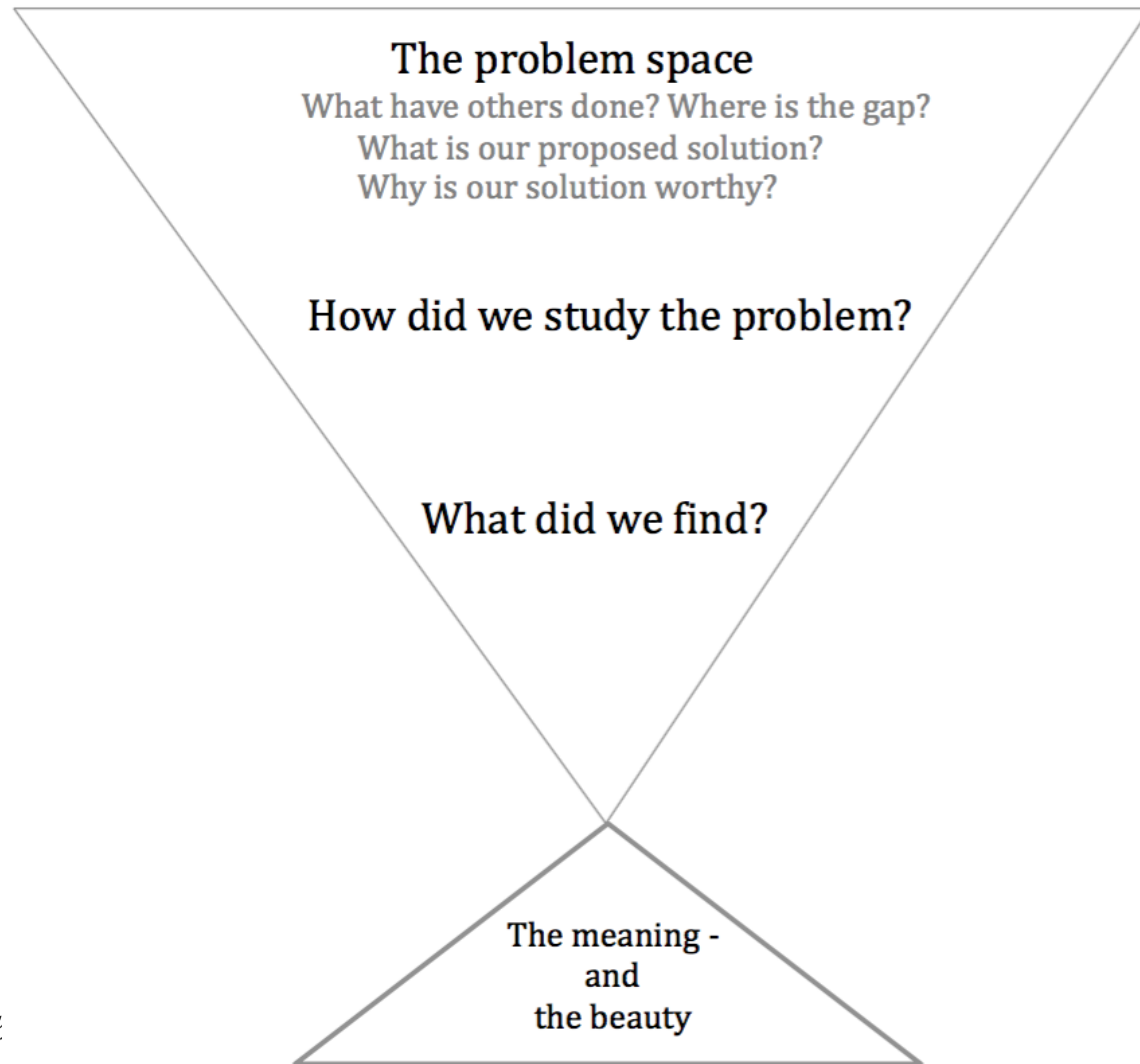
# Paper Structure

- The most common model:
  - Abstract
  - Introduction
  - Related Work
  - Your idea
  - Your implementation and experimentation/Methodology
  - Evaluation/Data analysis/Results
  - Discussion and Future work
  - Conclusion

The most important thing is the flow of the argument, the sectioning is often **discipline-specific**



# Rhetorical Positioning



# Rhetorical Positioning

- **Introduction** plays a key role in paper positioning
  - **Move 1**: establish a research territory
    - The problem is interesting, relevant, unsolved
    - Review previous research
  - **Move 2**: establish a niche
    - Indicate a gap in the previous research
  - **Move 3**: occupy a niche



# Intro Move 1, Example

***“The increasing interest*** in high angle-of-attack aerodynamics ***has heightened the need*** for computational tools suitable to predict the flowfield and aerodynamic coefficients in this regime. ***Of particular interest and complexity are*** the symmetric and asymmetric separated vortex flows which develop about slender bodies as the angle of attack is increased.” [Almosnino, 1985]



Claiming centrality





# Intro Move 1, Related Work

- Be well aware of all the relevant work in the field

MobiCom2011 Review #329A

===== Paper weaknesses =====

Related work review is weak...

network MIMO has been intensively studied

DySPAN 2012 Review #68B

===== Comments for author =====

How does your channel allocation algorithm compare with prior work?

IEEE PerCom 2014

=====

Small related work section.

MobileHCI 2015

----- Submission 200, Review 5 -----

Further, it is essential to discuss this work with prior work clearly.



# Intro Move 1, Related Work

- Be well aware of all the relevant work in the field
  - Briefly describe it in the intro
  - Extensively in a dedicated “Related Work” section
- Choosing the right tense
  - Past
    - When referring to work that is in a broader area, or not a direct competitor to your research
  - Present perfect
    - When referring to work that is the immediate area
  - Present
    - The most recent work, direct competitor



# Intro Move 2

- **Indicate a gap** in previous research
  - Both in the intro, and in the related work section where you will juxtapose your work against all the bordering papers
- From **counter-claiming** (something is wrong) to **continuing a tradition** (adding something)
- Avoid too many negative verbs and adjectives, use contrasts instead



# Intro Move 2, Example

“Directed diffusion [3] is one of a representative class of data dissemination mechanisms, specifically designed for a general class of applications in sensor networks. Directed diffusion provides robust dissemination through the use of multi-path data forwarding, **but the correct reception of all data messages is not assured**. We observed that in the context of sensor networks, data that flows from sources to sinks is generally tolerable of loss. **On the other hand, however, data that flows from sinks to sources for the purpose of control or management (e.g., re-tasking sensors) is sensitive to message loss**”

...

“Reliable point-to-point, or more appropriately, multicast transport mechanisms are well understood in conventional IP-style communication networks, where nodes are identified by their end-points. **However, these schemes (e.g., TCP, XTP [4], SRM [5]) cannot be efficiently applied to sensor networks** mainly because of the unique communication challenges presented by wireless sensor networks, including the need to support cluster-based communications, wireless multi-hop forwarding, application-specific operations, and lack of clean layering for the purposes of optimization.” [Wan et al., 2002]



# Intro Move 3

- **Fill the gap** indicated in Move 2
  - Signalled by a reference to the current text
  - Use of the present tense
- **Summarise your achievements**
  - Bullet point list of contributions
  - State the highlights from the results section
- Wrap up with a statement of the value of your research
- Avoid a roadmap paragraph in short papers



# Intro Example

“To the best of our knowledge, no prior work has demonstrated PHY layer-based WiFi localization on off-the-shelf platforms. Zhang et. al. [11] used signal amplitudes and phases on USRP platforms to demonstrate *location distinction*. We note that location distinction detects when a node’s location has changed (e.g., for security purposes), but does not need to establish uniqueness for each location. Localization is naturally a far stricter problem, especially when the target is sub-meter accuracies. PinLoc makes an early effort towards this goal – the main contributions may be summarized as follows.

- **We target the problem of *spot localization* where success is defined as the ability to place a device within a 1m x 1m area, called spots.** We break away from RSSI based schemes and explore the feasibility of using detailed PHY layer information.
- **We utilize the per-subcarrier frequency response as features of a location, and rely on machine learning algorithms to classify a device to one of the trained spots.** We use off-the-shelf Intel 5300 cards; the entire system relies on existing WiFi deployments, and requires no special installation.
- **We evaluate PinLoc at varying accuracy standards, namely, discriminating between seats in a lab, chairs in a cafeteria, and adjacent paintings in a museum.** We observe consistent accuracies under mobile/dynamic environments, outperforming Horus [4], the most accurate RSSI based localization.” [Sen et al., 2012]



# Introduction

- Your introductions should answer the following:
  - Why this issue important?
  - Why the problem difficult?
  - Why this contribution is stellar?
  - Why this is the most appropriate solution?
- If appropriate, use a figure in the introduction
- Let's read an example!



# Methodology

- Should convince the reader that what you did or propose to do is indeed **feasible** and **repeatable**
- The amount of detail depends on the publication venue
  - A paper submitted to a workshop on deep learning on IoT devices does not need the explanation of what a CNN or an Arduino board is
- Any methodological novelties should be highlighted and justified
- Figures help a lot





# Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Ebiers, de Legur, de Fezensac, de Chambray et le journal intime de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mohilow et qui rejoignent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

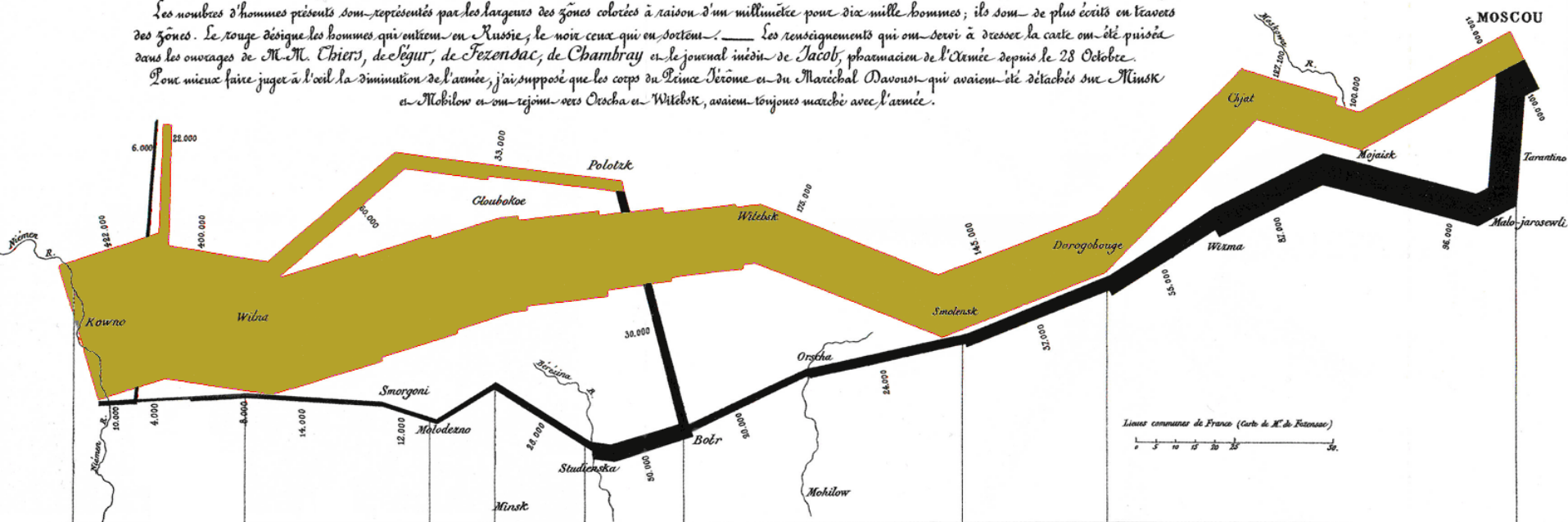
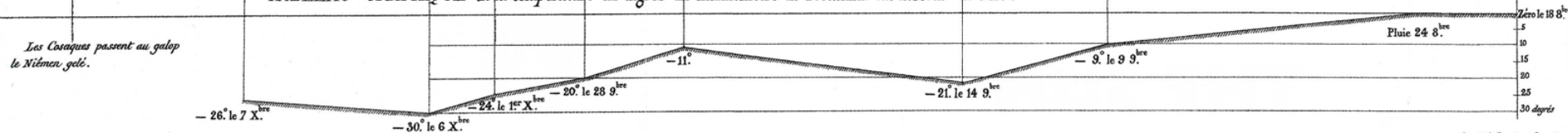


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Autog. par Regnier, R. Par. 5<sup>th</sup> Maria 5<sup>th</sup> 6<sup>th</sup> à Paris.

Imp. Lith. Regnier et Dorel.

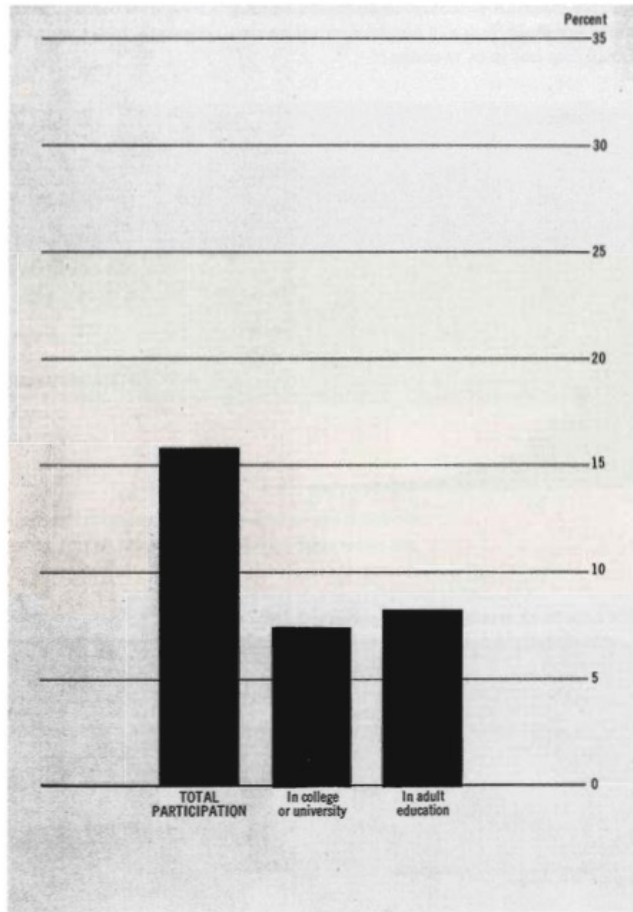


# Data Presentation

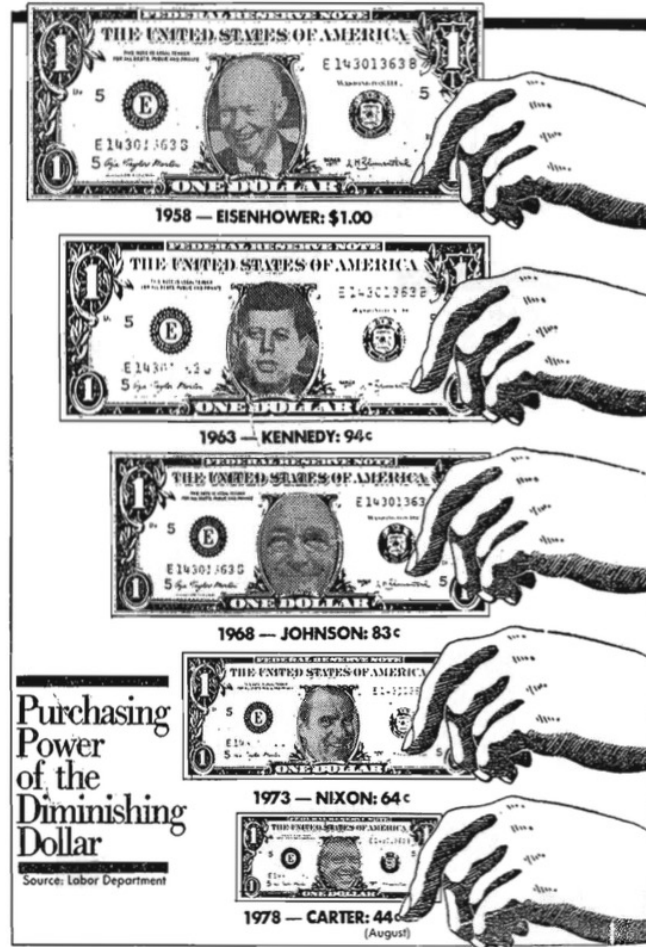
- Visual and verbal presentation of data should not be redundant, yet figures should not be left unexplained
- Visuals:
  - Aim for a high **data-ink ratio**
  - Maintain **data integrity**
  - Be rigorous: show error bars, label axes
  - Presentation clearly: your figures should be readable when printed on an A4-size paper



# Data Presentation Examples



Executive Office of the President, Office of Management and Budget, *Social Indicators, 1973* (Washington, D.C., 1973), p. 86.



*Washington Post*, October 25, 1978, p. 1.

# Data Presentation

- Data commentaries:
  - Use visualised data as a basis for your claim
  - Use both the main text, and figure captions
- Language focus:
  - Ensure proper referencing (Figure 4 shows...)
  - Toning the strength of your claim:
    - Likelihood
    - Distance
    - Combined qualifications

Pg 10 in Kayfetz's materials

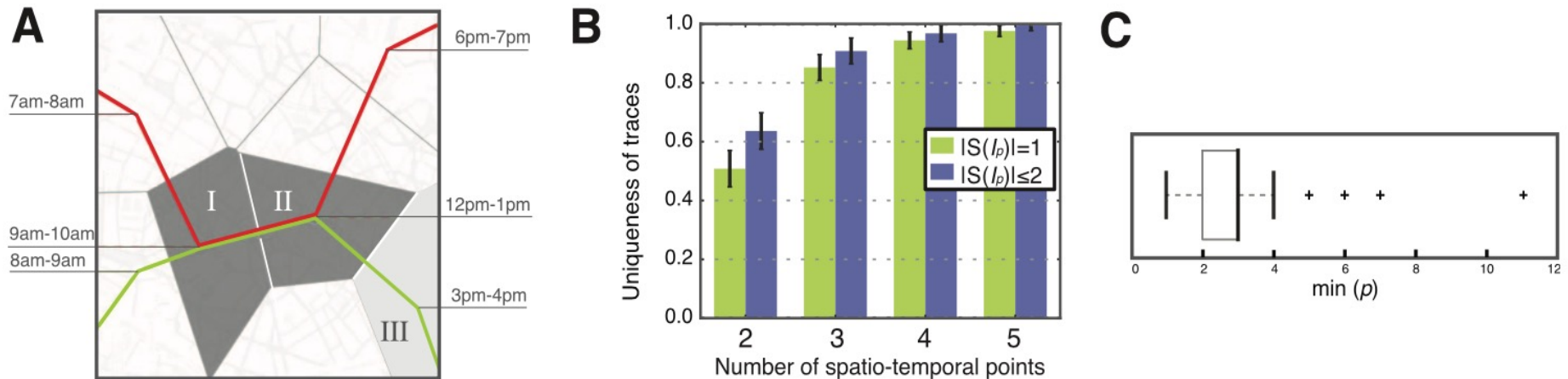


# Data Presentation

- Data commentaries:
  - **Why** are you showing a particular figure? Set the context for the reader
  - **Axes** should indicate (dependent) variables
  - **Lines** – one per each measurement group, explained, have a baseline where needed
  - **Trends** – do not assume they are obvious, guide the audience to view important points
  - **Exceptions** – do not hide exceptions, explain them
  - **Recap** – discuss why the results are significant, use to segue to the next figure



# Data Commentary Examples

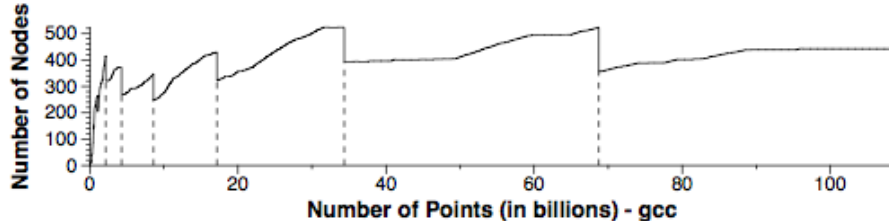


**Figure 2** | (A)  $I_{p=2}$  means that the information available to the attacker consist of two 7am-8am spatio-temporal points (I and II). In this case, the target was in zone I between 9am to 10am and in zone II between 12pm to 1pm. In this example, the traces of two anonymized users (red and green) are compatible with the constraints defined by  $I_{p=2}$ . The subset  $S(I_{p=2})$  contains more than one trace and is therefore not unique. However, the green trace would be uniquely characterized if a third point, zone III between 3pm and 4pm, is added ( $I_{p=3}$ ). (B) The uniqueness of traces with respect to the number  $p$  of given spatio-temporal points ( $I_p$ ). The green bars represent the fraction of unique traces, i.e.  $|S(I_p)| = 1$ . The blue bars represent the fraction of  $|S(I_p)| \leq 2$ . Therefore knowing as few as four spatio-temporal points taken at random ( $I_{p=4}$ ) is enough to uniquely characterize 95% of the traces amongst 1.5 M users. (C) Box-plot of the minimum number of spatio-temporal points needed to uniquely characterize every trace on the non-aggregated database. At most eleven points are enough to uniquely characterize all considered traces.



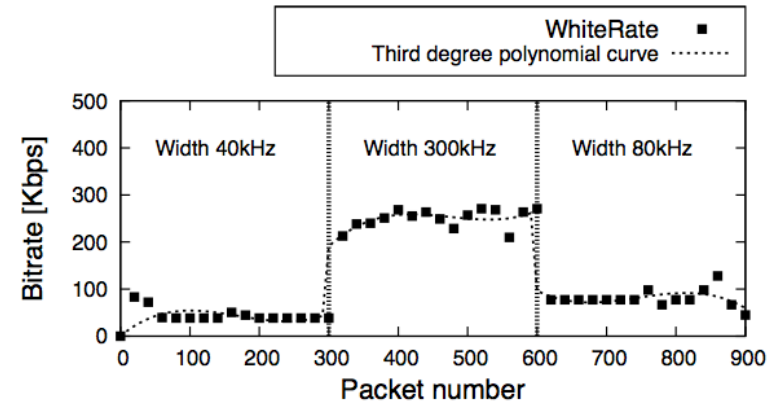


# Data Commentary Examples

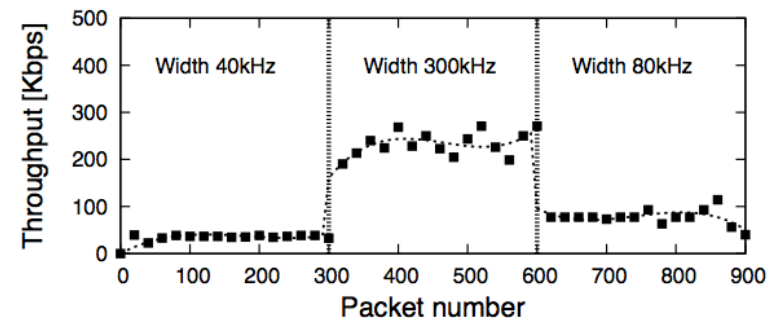


**Figure 6:** Number of nodes required to track the basic blocks of gcc with  $\epsilon = 10\%$ . While the number of nodes is far less than the worst case bounds that we estimated, the pattern of growth (due to splits) and rapid reduction (due to batched merging at points marked by dashed lines) can be clearly seen.

[Mysore et al, 2006]



(a) Bitrate.



(b) Throughput.

**Fig. 12. WhiteRate adapts to changing spectrum availability.** We modify available bandwidth from 40kHz to 300kHz and then to 80kHz.

[Pejovic and Belding, 2014]



# Discussion and Conclusion

- Showcase your main research point, and findings, state the potential impact
- Switch to past tense
- Common structure:
  - Background: research purpose, methodology
  - Summary of key results
  - Comment on key results
  - Limitations of the study
  - Recommendations for future research





# Discussion and Conclusion

## Example

“Congestion in an IEEE 802.11 network causes a drastic reduction in network performance. Critical to tackling this problem is the ability to identify and measure congestion. In this paper we have presented a passive technique that measures the utilization of the wireless medium in real-time. We then used the congestion measurement technique to develop a rate adaptation scheme called WOOF. Performance evaluation showed up to a three-fold gain in throughput with the use of WOOF in a congested network. In addition to WOOF, we believe that our congestion measurement technique can be used to design new solutions that perform well under congestion scenarios. For example, the CBT metric can be used for bandwidth estimation to facilitate effective flow admission control.” [Acharya et al, 2008]



# Abstract

- **Purpose:** invites the reader to your paper
  - Essential for review bidding, summary articles, cross disciplinary knowledge transfer
- **Length:**
  - As short as possible while still revealing what your paper is about, and what are the main findings
  - An average computer science publication abstract is 9.6 sentences, 232 words long



## Scholar

About 18,200 results (0.12 sec)

## Articles

Case law

My library

## Any time

Since 2015

Since 2014

Since 2011

Custom range...

## Sort by relevance

Sort by date

☒ include patents☒ include citations☒ Create alert**Mobility prediction-based smartphone energy optimization for everyday location monitoring**[Y Chon](#), [E Talipov](#), [H Shin](#), [H Cha](#) - ... of the 9th ACM conference on ..., 2011 - [dl.acm.org](#)

Abstract Monitoring a user's **mobility** during daily life is an essential requirement in providing advanced mobile services. While extensive attempts have been made to monitor user **mobility**, previous work has rarely addressed issues with battery lifetime in real ...

[Cited by 70](#) [Related articles](#) [All 5 versions](#) [Cite](#) [Save](#) [More](#)**Contextual conditional models for smartphone-based human mobility prediction**[TMT Do](#), [D Gatica-Perez](#) - Proceedings of the 2012 ACM conference on ..., 2012 - [dl.acm.org](#)

Abstract Human behavior is often complex and context-dependent. This paper presents a general technique to exploit this "multidimensional" contextual variable for human **mobility prediction**. We use an ensemble method, in which we extract different **mobility** patterns ...

[Cited by 29](#) [Related articles](#) [All 8 versions](#) [Cite](#) [Save](#) [More](#)**Smartphone-based applications for investigating falls and mobility**[C Tacconi](#), [S Mellone](#), [L Chiari](#) - ... Computing Technologies for ..., 2011 - [ieeexplore.ieee.org](#)

... iTUG has proven to be sensitive to age-related changes and fall risk **prediction** [12]. ... Systematic review of accuracy of screening instruments for **predicting** fall risk among independently living older ... A, Lundin-Olsson L. Changes in step-width during dual-task walking **predicts** falls ...

[Cited by 36](#) [Related articles](#) [Cite](#) [Save](#)**Evaluating mobility models for temporal prediction with high-granularity mobility data**[Y Chon](#), [H Shin](#), [E Talipov](#), [H Cha](#) - Pervasive computing and ..., 2012 - [ieeexplore.ieee.org](#)

... context of previously visited locations excessively filters out information necessary for **predicting** stay-duration. ... We use **mobility** monitoring application [8] to estimate d. To **predict** a user's behavior ... The accuracy of location **prediction** fa is the ratio of **predicted** probability of actual ...

[Cited by 33](#) [Related articles](#) [All 3 versions](#) [Cite](#) [Save](#) [More](#)**Where and what: Using smartphones to predict next locations and applications in daily life**[TMT Do](#), [D Gatica-Perez](#) - Pervasive and Mobile Computing, 2014 - Elsevier

... More precisely, we **predict** the next location of a user and which application he ... This approach allows modeling the interplay between the **predicted** variables to study relationships between ... context which allows us to learn useful generic models for **predicting** several dimensions ...

[Cited by 16](#) [Related articles](#) [All 8 versions](#) [Web of Science: 2](#) [Cite](#) [Save](#) [More](#)

# Abstract

- Answer the following:
  - What is your study about?
  - How was it done?
  - What was discovered?
  - What do the findings mean?
- Examples of opening sentences

Pg 18 in Kayfetz's  
materials



# Abstract Example

“The IEEE 802.11n standard defines channel bonding that allows wireless devices to operate on 40MHz channels by doubling their bandwidth from standard 20MHz channels. Increasing channel width increases capacity, but it comes at the cost of decreased transmission range and greater susceptibility to interference. However, with the incorporation of Multiple-Input Multiple-Output (MIMO) technology in 802.11n, devices can now exploit the increased transmission rates from wider channels with minimal sacrifice to signal quality and range. The goal of our work is to identify the network factors that influence the performance of channel bonding in 802.11n networks and make intelligent channel bonding decisions. We discover that channel width selection should consider not only a link’s signal quality, but also the strength of neighboring links, their physical rates, and interferer load. We use our findings to design and implement a network detector that successfully identifies interference conditions that affect channel bonding decisions in 100% of our test cases. Our detector can form the foundation for more robust and accurate algorithms that can adapt bandwidth to variations in channel conditions. Our findings allows us to predict the impact of network conditions on performance and make channel bonding decisions that maximize throughput.” [Deek et al, 2014]



# Abstract MadLibs!!

This paper presents a \_\_\_\_\_ method for \_\_\_\_\_  
(synonym for new) (sciencey verb)  
the \_\_\_\_\_. Using \_\_\_\_\_, the  
(noun few people have heard of) (something you didn't invent)  
\_\_\_\_\_ was measured to be \_\_\_\_\_ +/- \_\_\_\_\_  
(property) (number) (number)  
\_\_\_\_\_. Results show \_\_\_\_\_ agreement with  
(units) (sexy adjective)  
theoretical predictions and significant improvement over  
previous efforts by \_\_\_\_\_, et al. The work presented  
(Loser)  
here has profound implications for future studies of  
\_\_\_\_\_ and may one day help solve the problem of  
(buzzword)  
\_\_\_\_\_.  
(supreme sociological concern)

Keywords: \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_  
(buzzword) (buzzword) (buzzword)

JORGE CHAM © 2009



# Grammar and Style

- Academic writing is formal:
  - Researchers *looked* at the way strain *builds up* around a fault
  - Researchers *observed* the way strain *accumulates* around a fault.
  - Avoid contractions (won't == will not)
  - Avoid repetition (search for synonyms online)
  - Be consistent with: verb tenses, listing nouns/verbs, capitalisation of section headings, figure titles, etc.

Sloppy writing ->  
sloppy work



# How to Write a Paragraph

- A perfect paragraph
  - Topic sentence
  - Concrete Detail #1
  - Commentary
  - Transition
  - Concrete Detail #2
  - Commentary
  - Concluding Statement
- One idea per paragraph





# How to Write a Paragraph

“The goal of the preprocessing step is to dampen the breathing signal and improve the signal-to-interference-and-noise ratio (SINR) of the heartbeat signal. Recall that the phase of the RF signal is proportional to the composite displacement due to the inhale-exhale process and the pulsing effect. Since displacements due to the inhale-exhale process are orders of magnitude larger than minute vibrations due to heartbeats, the RF phase signal is dominated by breathing. However, the acceleration of breathing is smaller than that of heartbeats. This is because breathing is usually slow and steady while a heartbeat involves rapid contraction of the muscles. Thus, we can dampen breathing and emphasize the heartbeats by operating on a signal proportional to acceleration as opposed to displacement.” [Zhao et al., 2015]



# How to Write a Paragraph

- Tips for writing a good paragraph
  - First sentence summarizes a paragraph
  - Last sentence segues into the next paragraph
  - Short paragraphs indicate that you did not explain the idea in sufficient detail
  - Long paragraphs indicate that you are either too verbose or have more than one idea per paragraph
  - Be as concise as possible, avoid filler words and adjectives



# How to Write a Paragraph



# Academic Writing for Foreign Students

- Common pitfalls  
(observed on a very modest sample)
  - Too much passive voice
  - Frequent use of “-ing” form, use present simple tense or nouns instead
  - Long, complex sentences, break into shorter sentences
  - More than one idea per sentence
  - Redundant, filler words (“in order to”, “note that”)
  - Not using a spellchecker



# SCIENCE ARTICLES: A GUIDE

	AVERAGE SENTENCE IS EASY TO UNDERSTAND	AVERAGE SENTENCE IS HARD TO UNDERSTAND
SUBJECT MATTER IS COMPLEX	GREAT WRITING	TYPICAL WRITING
SUBJECT MATTER IS SIMPLE	HONEST WRITING	PROBABLY JUST BULLSHIT

smbc-comics.com



# Techniques to Improve Your Writing

- Read your writing aloud: get the right rhythm, hear if a sentence is too long/short
- Check if you are still on the track, ask:
  - Why am I saying this?
  - Does it clarify the point?
  - Does an example support this idea?
- Write for the reader, not for yourself!
- Don't get attached to your writing!
  - Default mindset: the first version is bad



Let's practice!

# References

- Essential books:
  - Strunk, W. , White, E. B. (1972). *The elements of style*. Macmillan.
  - Tufle, E. R. (1983) *The visual display of quantitative information*. CT Graphics, Cheshire.



# References

- Sen, Souvik, et al. "You are facing the mona lisa: spot localization using phy layer information." ACM MobiSys, 2012.
- Deek, Lara, et al. "Intelligent channel bonding in 802.11 n WLANs." IEEE Transactions on Mobile Computing, 13.6 (2014): 1242-1255.
- Pejovic, Veljko, and Elizabeth M. Belding. "WhiteRate: A Context-Aware Approach to Wireless Rate Adaptation. IEEE Transactions on " Mobile Computing, 13.4 (2014): 921-934.
- Acharya, Prashanth Aravinda Kumar, et al. "Congestion-aware rate adaptation in wireless networks: A measurement-driven approach." IEEE SECON, 2008.
- Wan, Chieh-Yih, et al. "PSFQ: a reliable transport protocol for wireless sensor networks." ACM international workshop on Wireless sensor networks and applications, 2002
- de Montjoye, Yves-Alexandre, et al. "Unique in the Crowd: The privacy bounds of human mobility." Scientific reports 3 (2013).
- Mysore, Shashidhar, et al. "Profiling over adaptive ranges." IEEE CGO 2006





# Thank you

Dr Veljko Pejović  
Veljko.Pejovic@fri.uni-lj.si



University of Ljubljana  
Faculty of Computer and  
Information Science