The UMAP Journal

Vol. 21, No. 3

Subscription Rates for 2000 Calendar Year: Volume 21

Membership Plus for Individual Subscribers

Individuals subscribe to *The UMAP Journal* through COMAP's Membership Plus. This subscription includes print copies of quarterly issues of *The UMAP Journal*, our annual collection *UMAP Modules: Tools for Teaching*, our organizational newsletter *Consortium*, on-line membership that allows members to search our on-line catalog, download COMAP print materials, and reproduce for use in their classes, and a 10% discount on all COMAP materials.

(Domestic)	**#2020**	**$69**
(Outside U.S.)	**#2021**	**$79**

Institutional Plus Membership Subscribers

Institutions can subscribe to the *Journal* through either Institutional Pus Membership, Regular Institutional Membership, or a Library Subscription. Institutional Plus Members receive two print copies of each of the quarterly issues of *The UMAP Journal*, our annual collection *UMAP Modules: Tools for Teaching*, our organizational newsletter *Consortium*, on-line membership that allows members to search our on-line catalog, download COMAP print materials, and reproduce for use in **any class taught in the institution, and** a 10% discount on all COMAP materials.

(Domestic)	**#2070**	**$395**
(Outside U.S.)	**#2071**	**$415**

Institutional Membership Subscribers

Regular Institutional members receive only print copies of *The UMAP Journal*, our annual collection *UMAP Modules: Tools for Teaching*, our organizational newsletter *Consortium*, **and** a 10% discount on all COMAP materials.

(Domestic)	**#2040**	**$165**
(Outside U.S.)	**#2041**	**$185**

Library Subscriptions

The Library Subscription includes quarterly issues of *The UMAP Journal* and our annual collection *UMAP Modules: Tools for Teaching* and our organizational newsletter *Consortium*.

(Domestic)	**#2030**	**$140**
(Outside U.S.)	**#2031**	**$160**

To order, send a check or money order to COMAP, or call toll-free
1-800-77-COMAP (1-800-772-6627).

The UMAP Journal is published quarterly by the Consortium for Mathematics and Its Applications (COMAP), Inc., Suite 210, 57 Bedford Street, Lexington, MA, 02420, in cooperation with the American Mathematical Association of Two-Year Colleges (AMATYC), the Mathematical Association of America (MAA), the National Council of Teachers of Mathematics (NCTM), the American Statistical Association (ASA), the Society for Industrial and Applied Mathematics (SIAM), and The Institute for Operations Research and the Management Sciences (INFORMS). The Journal acquaints readers with a wide variety of professional applications of the mathematical sciences and provides a forum for the discussion of new directions in mathematical education (ISSN 0197-3622).

Second-class postage paid at Boston, MA
and at additional mailing offices.
Send address changes to:
The UMAP Journal
COMAP, Inc.
57 Bedford Street, Suite 210, Lexington, MA 02420

Vol. 21, No. 3 2000

Table of Contents

Interdisciplinary Contest in Modeling (ICM)

The Elephant Population Problem

Publisher's Editorial

COMAP on the Web

Solomon A. Garfunkel
Executive Director
COMAP, Inc.
57 Bedford St., Suite 210
Lexington, MA 02420
s.garfunkel@mail.comap.com

As you may be aware, COMAP has over the past year updated our Web site. Not only are there clearer and more detailed descriptions of our products and projects, but we have initiated a Web membership and made our major publications available online. In addition, this year we will move (almost) exclusively to Web registration and problem delivery for the Mathematical Contest in Modeling (MCM) and the High School Contest in Modeling (HiMCM). All of our new supplementary materials are being designed with the Web clearly in mind. We understand that the Internet is increasingly the delivery method of choice for educational materials.

Moreover, we are planning a number of new projects that use the Web as a vehicle for the delivery of distance learning. It has been a while since we created television courses as a way of promoting lifelong learning. As we have developed new curricula at the secondary level, we have come to understand the importance of teacher preparation and enhancement. New teachers need to be prepared to face the challenges of the content and pedagogy of Standards-based curricula, as well as new developments in technology. Web-based courses have become increasingly important in the continuing education of teachers and students. We see this as an important area for COMAP's future.

This year has seen the publication of our fourth course in the *Mathematics: Modeling Our World* series. We are currently working on college level texts in Precalculus and in College Algebra, all with W.H. Freeman as publisher. And we look forward to the publication next year, with Brooks-Cole, of a new text in mathematical methods for secondary school teachers, embracing a modeling- and applications-based approach. We expect and intend to be in the business of textbook development for many many years to come. But we are aware that a large part of that future will be in electronic publishing—and we are investing in that future.

The UMAP Journal 21 (3) (2000) 211–212.

In Memoriam: Ross L. Finney

We write here to mourn the passing of Ross Finney this August.

I am sure that most of the readers of *The UMAP Journal* know Ross from his work on the last several editions of the Thomas calculus texts. As important as that work has been, Ross Finney was also a pioneer in mathematics education in the pre-COMAP era. In fact, Ross was a co-principal investigator on the original UMAP grant from NSF that led directly to the founding of COMAP.

We remember Ross as the founding editor of *The UMAP Journal*; he continued that effort through the first five years of its existence. It is fair to say that the *Journal* that you hold in your hands would likely never have existed without his efforts. We here at COMAP, and readers of the *Journal* throughout the world, owe him a large debt of gratitude. He will be sorely missed.

Obituaries

Ross Lee Finney III; math teacher wrote calculus textbooks. 2000. *Los Angeles Times* (18 August 2000): B-6. `http://www.latimes.com/print/metro/20000818/t000077771.html` .

Saxon, Wolfgang. 2000. Ross Lee Finney III, 67, author of widely used math textbooks. *New York Times* (16 August 2000). `http://www10.nytimes.com/yr/mo/day/news/national/obit-r-finney.html` .

About the Author

Sol Garfunkel received his Ph.D. in mathematical logic from the University of Wisconsin in 1967. He was at Cornell University and at the University of Connecticut at Storrs for eleven years and has dedicated the last 20 years to research and development efforts in mathematics education. He has been the Executive Director of COMAP since its inception in 1980.

He has directed a wide variety of projects, including UMAP (Undergraduate Mathematics and Its Applications Project), which led to the founding of this *Journal*, and HiMAP (High School Mathematics and Its Applications Project), both funded by the NSF. For Annenberg/CPB, he directed three telecourse projects: *For All Practical Purposes* (in which he appeared as the on-camera host), *Against All Odds: Inside Statistics*, and *In Simplest Terms: College Algebra*. He is currently co-director of the Applications Reform in Secondary Education (ARISE) project, a comprehensive curriculum development project for secondary school mathematics.

Modeling Forum

Results of the 2000 Mathematical Contest in Modeling and Interdisciplinary Contest in Modeling

Frank Giordano, MCM Director
COMAP, Inc.
57 Bedford St., Suite 210
Lexington, MA 02420
f.giordano@mail.comap.com

Chris Arney, ICM Co-Director
U.S. Military Academy
West Point, NY
ad6819@exmail.usma.army.mil

John H. Grubbs, ICM Co-Director
Tulane University
New Orleans, LA

Introduction

COMAP is pleased to announce the results of the 16th annual Mathematical Contest in Modeling (MCM) and the 2nd annual Interdisciplinary Contest in Modeling (ICM). This year 495 teams representing 231 institutions from 9 countries spent the first weekend in February working on applied mathematics and interdisciplinary problems.

The 2000 MCM/ICM began at 12:01 A.M. on Friday, Feb. 4 and officially ended at 5:00 P.M. on Monday, Feb. 7, 2000 (local time). Teams of two or three undergraduates were to research and submit an optimal solution for one of

The UMAP Journal 21 (3) (2000) 213–239.

three open-ended modeling problems. After a weekend of hard work, typed solution papers were mailed to COMAP. Twelve of the top papers appear in this issue of *The UMAP Journal*.

This year's Problem A was dedicated to the memory of Dr. Robert Machol, former chief scientist of the Federal Aviation Agency. Dr. Machol posed the problem when the FAA was considering adding software to the air traffic control system that would alert controllers to potential problems and thus improve safety and reduce workload. In addition to solving the problem, participants were asked to write a summary that could be presented to the FAA Administrator, Ms. Jane Garvey.

Problem B this year sought to model the assignment of radio channels to a symmetric network of transmitter locations so as to avoid interference. Many groups came to the conclusion that the pure mathematical solution needed some practical applications and included those as well.

This year's ICM Problem C offered information regarding the need to keep the elephant population in a national park in South Africa down to 11,000 while avoiding having to destroy any of the animals. A contraceptive dart has been developed that prevents conception for two years. Participants were to investigate a strategy for how to use the dart successfully. Six specific questions were posed and data were offered about emigration patterns, gender ratios, elephant conception patterns, and new calf survival rates.

Results and winning papers from the first fifteen contests were published in special issues of *Mathematical Modeling* (1985–1987) and *The UMAP Journal* (1985–1999). The 1994 volume of *Tools for Teaching*, commemorating the tenth anniversary of the contest, contains the 20 problems used in the first ten years of the contest and a winning paper for each. Limited quantities of that volume and of the special MCM issues of the *Journal* for the last few years are available from COMAP.

Problem A: Air Traffic Control

Dedicated to the memory of Dr. Robert Machol,
former chief scientist of the Federal Aviation Agency

To improve safety and reduce air traffic controller workload, the Federal Aviation Agency (FAA) is considering adding software to the air traffic control system that would automatically detect potential aircraft flight path conflicts and alert the controller. To that end, an analyst at the FAA has posed the following problems.

- Requirement A: Given two airplanes flying in space, when should the air traffic controller consider the objects to be too close and to require intervention?

- Requirement B: An airspace sector is the section of three-dimensional airspace that one air traffic controller controls. Given any airspace sector, how do we measure how complex it is from an air traffic workload perspective? To what extent is complexity determined by the number of aircraft simultaneously passing through that sector

- at any one instant?
- during any given interval of time?
- during a particular time of day?

How does the number of potential conflicts arising during those periods affect complexity? Does the presence of additional software tools to automatically predict conflicts and alert the controller reduce or add to this complexity?

In addition to the guidelines for your report, write a summary (no more than two pages) that the FAA analyst can present to Jane Garvey, the FAA Administrator, to defend your conclusions.

Problem B: Radio Channel Assignments

We seek to model the assignment of radio channels to a symmetric network of transmitter locations over a large planar area, so as to avoid interference. One basic approach is to partition the region into regular hexagons in a grid (honeycomb-style), as shown in **Figure 1**, where a transmitter is located at the center of each hexagon.

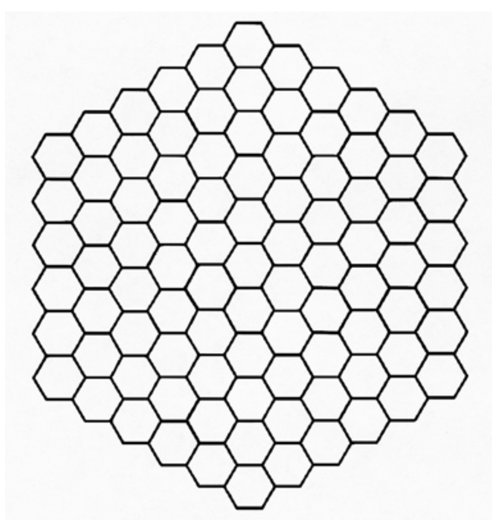


Figure 1. Honeycomb grid of hexagons.

An interval of the frequency spectrum is to be allotted for transmitter frequencies. The interval will be divided into regularly spaced channels, which we represent by integers 1, 2, 3,. . . . Each transmitter will be assigned one positive integer channel. The same channel can be used at many locations, provided that interference from nearby transmitters is avoided.

Our goal is to minimize the width of the interval in the frequency spectrum that is needed to assign channels subject to some constraints. This is achieved with the concept of a span. The span is the minimum, over all assignments satisfying the constraints, of the largest channel used at any location. It is not required that every channel smaller than the span be used in an assignment that attains the span.

Let s be the length of a side of one of the hexagons. We concentrate on the case that there are two levels of interference.

- Requirement A: There are several constraints on frequency assignments:
  - No two transmitters within distance $4s$ of each other can be given the same channel.
  - Due to spectral spreading, transmitters within distance $2s$ of each other must not be given the same or adjacent channels: Their channels must differ by at least 2.

  Under these constraints, what can we say about the span in Figure 1?
- Requirement B: Repeat Requirement A, assuming the grid in the example spreads arbitrarily far in all directions.
- Requirement C: Repeat Requirements A and B, except assume now more generally that channels for transmitters within distance $2s$ differ by at least some given integer k, while those at distance at most $4s$ must still differ by at least one. What can we say about the span and about efficient strategies for designing assignments, as a function of k?
- Requirement D: Consider generalizations of the problem, such as several levels of interference or irregular transmitter placements. What other factors may be important to consider?
- Requirement E: Write an article (no more than 2 pages) for the local newspaper explaining your findings.

ICM Problem: Elephants: When is Enough, Enough?

> “Ultimately, if a habitat is undesirably changed by elephants, then their removal should be considered—even by culling.”
>
> *National Geographic* (Earth Almanac) (December 1999)

A large national park in South Africa contains approximately 11,000 elephants. Management policy requires a healthy environment that can maintain a stable herd of 11,000 elephants. Each year, park rangers count the elephant population. During the past 20 years whole herds have been removed to keep the population as close to 11,000 as possible. This process involved shooting (for the most part) and occasionally relocating approximately 600 to 800 elephants per year.

Recently, there has been a public outcry against the shooting of these elephants. In addition, it is no longer feasible to relocate even a small population of elephants each year. A contraceptive dart, however, has been developed that can prevent a mature elephant cow from conceiving for a period of two years.

Here is some information about the elephants in the park:

- There is very little emigration or immigration of elephants.
- The gender ratio is very close to 1:1 and control measures have endeavored to maintain parity.
- The gender ratio of newborn calves is also about 1:1. Twins are born about 1.35% of the time.
- Cows first conceive between the ages of 10 and 12 and produce, on average, a calf every 3.5 years until they reach an age of about 60. Gestation is approximately 22 months.
- The contraceptive dart causes an elephant cow to come into oestrus every month (but not conceiving). Elephants usually have courtship only once in 3.5 years, so the monthly cycle can cause additional stress.
- A cow can be darted every year without additional detrimental effects. A mature elephant cow will not be able to conceive for 2 years after the last darting.
- Between 70% and 80% of newborn calves survive to age 1 year. Thereafter, the survival rate is uniform across all ages and is very high (over 95%), until about age 60; it is a good assumption that elephants die before reaching age 70.
- There is no hunting and negligible poaching in the park.

The park management has a rough data file of the approximate ages and gender of the elephants that they have transported out of the region during the past two years. These data are available on the Web: `www.comap.com/icm/icm2000data.xls`. Unfortunately, no data are available for the elephants that were shot or that remain in the park.

Your overall task is to develop and use models to investigate how the contraceptive dart might be used for population control. Specifically:

- Task 1: Develop and use a model to speculate about the likely survival rate for elephants aged 2 to 60. Also speculate about the current age structure of the elephant population.
- Task 2: Estimate how many cows would need to be darted each year to keep the population fixed at approximately 11,000 elephants. Show how the uncertainty in the data at your disposal affects your estimate. Comment on any changes in the age structure of the population and how this might affect tourists. (You may want to look ahead about 30–60 years.)
- Task 3: If it were feasible to relocate between 50 and 300 elephants per year, how would this reduce the number of elephants to be darted? Comment on the trade-off between darting and relocation.
- Task 4: Some opponents of darting argue that if there were a sudden loss of a large number of elephants (due to disease or uncontrolled poaching), even if darting stopped immediately, the ability of the population to grow again would be seriously impeded. Investigate and respond to this concern.
- Task 5: The management in the park is skeptical about modeling. In particular, they argue that a lack of complete data makes a mockery of any attempt to use models to guide their decisions. In addition to your technical report, include a carefully crafted report (3-page maximum) written explicitly for the park management that responds to their concerns and provides advice. Also, suggest ways to increase the park managers' confidence in your model and in your conclusions.
- Task 6: If your model works, other elephant parks in Africa would be interested in using it. Prepare a darting plan for parks of various sizes (300–25,000 elephants), with slightly different survival rates and transportation possibilities.

The Results

The solution papers were coded at COMAP headquarters so that names and affiliations of the authors would be unknown to the judges. Each paper was read preliminarily by two "triage" judges at Southern Connecticut State University (Problem A), Carroll College (Montana) (Problem B), or University of New Hampshire (Problem C). At the triage stage, the summary and overall organization were important. If the judges' scores diverged for a paper, the judges conferred; if they still did not agree, a third judge evaluated the paper.

Final judging took place at Harvey Mudd College, Claremont, California. The judges classified the papers as follows:

The twelve papers that the judges designated as Outstanding appear in this special issue of *The UMAP Journal*, together with commentaries. We list

	Outstanding	Meritorious	Honorable Mention	Successful Participation	Total
Air Traffic Control	4	21	45	84	154
Channel Assignment	5	43	72	151	271
Elephant Population	3	12	18	37	70
	12	76	135	272	495

those teams and the Meritorious teams (and advisors) below; the list of all participating schools, advisors, and results is in the **Appendix**.

Outstanding Teams

Institution and Advisor	**Team Members**

Air Traffic Control Papers

"Air Traffic Control"	
Duke University	Samuel Westmoreland Malone
Durham, NC	Jeffrey Abraham Mermin
David P. Kraines	Daniel Bertrand Neill
"The Safe Distance Between Airplanes and the Complexity of an Airspace"	
Governor's School	Finale Doshi
Richmond, VA	Rebecca Lessem
Crista Hamilton	David Mooney
"The Iron Laws of Air Traffic Control"	
U.S. Military Academy	Kevin Arnett
West Point, NY	Jonathan S. Gibbs
David Bailey	John J. Horton
"You Make the Call: Feasibility of Computerized Aircraft Control"	
University of Colorado	Richard D. Younger
Boulder, CO	Martin B. Linck
Anne M. Dougherty	William P. Woesner

Channel Assignment Papers

"A Channel Assignment Model:
The Span Without a Face"
California Polytechnic State University
San Luis Obispo, CA
Thomas O'Neil

Jeffrey Mintz
Aaron Newcomer
James Price

"'We're Sorry, You're Outside the Coverage Area'"
Lewis and Clark College
Portland, OR
Robert W. Owens

Robert E. Broadhurst
William J. Shanahan
Michael D. Steffen

"Utilize the Limited Frequency Resources
Efficiently"
National University of Defence Technology
Changsha, Hunan, China
Wu Meng Da

Chu Rui
Xiu Baoxin
Zong Ruidi

"Groovin' with the Big Band(width)"
Wake Forest University
Winston–Salem, NC
Edward Allen

Daniel J. Durand
Jacob M. Kline
Kevin M. Woods

"Radio Channel Assignments"
Washington University
St. Louis, MO
Hiro Mukai

Justin Goodwin
Dan Johnston
Adam Marcus

Elephant Population Papers

"Elephant Population: A Linear Model"
Harvey Mudd College
Claremont, CA
Michael Moody

Nathan Cappallo
Daniel Osborn
Timothy Prescott

"A Computational Solution for Elephant
Overpopulation"
North Carolina School of Science and Mathematics
Durham, NC
Dot Doyle and Dan Teague

Jesse Crossen
Aaron Hertz
Danny Morano

"EigenElephants: When Is Enough, Enough?"
North Carolina School of Science and Mathematics
Durham, NC
Dot Doyle and Dan Teague

David Marks
Jim Sukha
Anand Thakker

Meritorious Teams

Air Traffic Control Papers (21 teams)
Chongqing University, Chongqing, China (Chen Yihua)
Drake University, Des Moines, IA (Alexander F. Kleiner)
East China Univ. of Science & Technology, Shanghai, China (Lu Yuanhong)
First Middle School of Jiading, Shanghai, China (Fang Yunping)
Harvey Mudd College, Claremont, CA (Zachary Dodds)
Lafayette College, Easton, PA (Thomas Hill)
Peking University, Beijing, China (Deng Minghua)
Rose-Hulman Institute of Technology, Terre Haute, IN (Frank Young)
Science School of Xi'an Jiaotong University, Xi'an, Shaanxi, China (He Xiaoliang)
Simpson College, Indianola, IA (M.E. Waggoner)
Stetson University, Deland, FL (Lisa O. Coulter)
Trinity University, San Antonio, TX (Tarynn Witten)
University College Cork, Cork, Ireland (Michael Quinlan)
University of Alaska Fairbanks, Fairbanks, AK (Chris Hartman)
University of Cincinnati, Cincinnati, OH (Charles Groetsch)
Univ. of Colorado at Colorado Springs, Colorado Springs, CO (Jon Epperson)
University of Saskatchewan, Saskatoon, SK, Canada (Raj Srinivasan)
University of Science & Technology of China, Hefei, Anhui, China (Sun Liang)
Worcester Polytechnic Institute, Worcester, MA (Bogdan Vernescu)
Youngstown State University, Youngstown, OH (Steve Hanzely)
Youngstown State University, Youngstown, OH (Thomas Smotter)
Zhejiang University, Hangzhou, Zhejiang, China (Yang Qifan)
Zhejiang University, Hangzhou, Zhejiang, China (He Yong)

Channel Assignment Papers (43 teams)
Anhui University, Hefei, Anhui, China (Chen Junsheng)
Asbury College, Wilmore, KY (Kenneth P. Rietz)
Beijing University of Post & Telecomm, Beijing, Beijing, China (He Zuguo)
Beijing University of Post & Telecomm, Beijing, Beijing, China
(Sun Hongxiang)
Calvin College, Grand Rapids, MI (Dorothea Pronk)
China University of Mining & Technology, Xuzhou, Jiangsu, China
(Wu Zongxiang)
East China Univ. of Science & Technology, Shanghai, China (Shi Jinsong)
East China Univ. of Science & Technology, Shanghai, China (Xiwen Lu)
Fudan University, Shanghai, China (Cai Zhijie)
Gettysburg College, Gettysburg, PA (James P. Fink)
Grinnell College, Grinnell, IA (Marc Chamberland)
Harbin Institute of Technology, Harbin, Heilongjiang, China (Wang Xuefeng)
Harvey Mudd College, Claremont, CA (Michael Moody)
Institution of Information Science & Engineering, Shenyang, Liaoning, China
(Xiao Wendong)
Luther College, Decorah, IA (Reginald D. Laursen)
MIT, Cambridge, MA (Michael Brenner)
Mt. Mercy College, Cedar Rapids, IA (Kent R. Knopp)

National U. of Defence Technology, Chang Sha, Hunan, China (Cheng Lizhi)
Northwestern Polytechnical University, Xi'an, Shaanxi, China (Liu Xiaodong)
Northwestern Polytechnical University, Xi'an, Shaanxi, China (Peng Guohua)
Pacific Lutheran University, Tacoma, WA (Rachid Benkhalti)
Paivola College, Tarttila, Finland (Esa Lappi)
Peking University, Beijing, China (Lei Gongyan)
Rose-Hulman Institute of Technology, Terre Haute, IN (David Rader)
Shanghai Foreign Languages School, Shanghai, China (Wan Baihe)
Shanghai Jiao Tong University, Shanghai, China (Zhou Gang)
South China Univ. of Technology, Guangzhou, Guangdong, China (Fu Hongzhuo)
Southeast University, Nanjing, China (Huang Jun) (two teams)
Trinity University, San Antonio, TX (Allen Holder)
Tsinghua University, Beijing, China (Hu Zhiming)
U.S. Military Academy, West Point, NY (Greg Parnell)
University of Alaska Fairbanks, Fairbanks, AK (Chris Hartman)
Univ. of Michigan—Dearborn, Dearborn, MI (David James)
University of New South Wales, Sydney, Australia, (James Franklin)
University of Richmond, Richmond, VA (Kathy W. Hoke)
University of Toronto, Toronto, Ontario, Canada (Nicholas A. Derzko) (two teams)
Wuhan Univ. of Hydraulic & Engineering, Wuban, Hubei, China (Peng Zuzeng)
Yale University, New Haven, CT (Steven Orszag)
Zhejiang University, Hangzhou, Zhejiang, China (Yang Qifan)
Zhejiang University, Hangzhou, Zhejiang, China (He Yong)

Elephant Population Papers (12 teams)
Bloomsburg University, Bloomsburg, PA (Kevin Ferland & Scott Inch)
California Academy of Math & Science, Carson, CA (Brian R. Lawler)
China University of Mining & Tech., Xuzhou, Jiangsu, China (Zhou Shengwu)
MIT, Cambridge, MA (Michael P. Brenner & Lakshmirarayanan Muhadevan)
Northwestern Polytechnical University, Xi'an, Shaanxi, China (Xu Wei & Wang Mingyu)
Peking University, Beijing, Beijing, China (Shao Min & Zhang Tao)
Southeast University, Nanjing, China (Chen Enshui)
U.S. Military Academy, West Point, NY (Michael Jaye & Greg Fleming)
Univ. of Science & Tech. of China, Hefei, Anhui, China (Wan Qian)
Youngstown State University, Youngstown, OH (Scott Martin)
Zhejiang University, Hangzhou, Zhejiang, China (Cong Zhang, Chu Jaiowu , & He Yong)
Zhejiang University, Hangzhou, Zhejiang, China (Cong Zhang, Chu Jaiowu, & Yang Qifan)

Awards and Contributions

Each participating MCM advisor and team member received a certificate signed by the Contest Director and the appropriate Head Judge.

INFORMS, the Institute for Operations Research and the Management Sciences, gave a cash award and a three-year membership to each member of the teams from the University of Colorado (Air Traffic Control Problem), Washington University (Channel Assignment Problem), and North Carolina School of Science and Mathematics (the team of Jesse Crossen, Aaron Hertz, and Danny Morano) (Elephant Population Problem). Moreover, INFORMS gave free one-year memberships to all members of Meritorious and Honorable Mention teams.

The Society for Industrial and Applied Mathematics (SIAM) designated as SIAM Winners the teams from U.S. Military Academy, West Point, NY (Air Traffic Control Problem) and Wake Forest University (Channel Assignment Problem). Each of the team members was awarded a $300 cash prize. Their schools were given framed certificates hand-lettered in gold leaf. Both teams presented their results at a special Minisymposium of the SIAM Annual Meeting in Puerto Rico in July.

The Mathematical Association of America (MAA) designated as MAA Winners the teams from the Duke University (Air Traffic Control) and California Polytechnic State University (Channel Assignment Problem). The team from California Polytechnic State University presented their solution at a special session of the MAA Mathfest in Los Angeles in August. Each team member was presented a certificate by MAA President Thomas Banchoff.

Judging

MCM

Director
Frank R. Giordano, COMAP, Lexington, MA

Associate Directors
Robert L. Borrelli, Mathematics Dept., Harvey Mudd College, Claremont, CA
William Fox, Chair, Dept. of Mathematics, Francis Marion University, Florence, SC

Air Traffic Control Problem

Head Judge
Martin Keener, Executive Vice President, Oklahoma State University, Stillwater, OK

Associate Judges
Ron Barnes, University of Houston—Downtown, Houston, TX (MAA)
Patrick Driscoll, Dept. of Mathematical Sciences, U.S. Military Academy, West Point, NY (INFORMS)

David L. Elliott, Institute for Systems Research, University of Maryland, College Park, MD (SIAM)
Gordon Erlebacher, Dept. of Computer Science and Information Technology, Florida State University, Tallahassee, FL
Richard Haberman, Mathematics Dept., Southern Methodist University, Dallas, TX (SIAM)
Mark Levinson, Edmonds, WA (SIAM)
Theresa M. Sandifer, Southern Connecticut State University, New Haven, CT (Triage)

Triage Judges
Head Triage Judge
Theresa M. Sandifer, Southern Connecticut State University, New Haven, CT
Associate Triage Judges
Ross Gingrich, Southern Connecticut State University
Cynthia B. Gubitose, Western Connecticut State University, Danbury, CT
Ronald E. Kutz, Western Connecticut State University, Danbury, CT
C. Edward Sandifer, Western Connecticut State University, Danbury, CT
Jim Wohlever, Western Connecticut State University, Danbury, CT

Channel Assignment Problem

Head Judge
Maynard Thompson, Mathematics Dept., University of Indiana, Bloomington, IN

Associate Judges
Paul Boisen, Defense Dept., Ft. Meade, MD
James Case, Baltimore, Maryland
Lisette de Pillis, Mathematics Dept., Harvey Mudd College, Claremont, CA
Doug Faires, Dept. of Mathematics and Statistics, Youngstown State University, Youngstown, OH
Jerry Griggs, Dept. of Mathematics, University of South Carolina, Columbia, SC (SIAM)
Jeff Hartzler, Dept. of Mathematics, Penn State University, Middletown, PA (MAA)
Mario Juncosa, RAND Corporation, Santa Monica, CA
Deborah Levinson, Dept. of Mathematics, Colorado College, Colorado Springs, CO
Veena Mendiratta, Lucent Technologies, Naperville, IL
Don Miller, Dept. of Mathematics, St. Mary's College, Notre Dame, IN
Mark Parker, Dept. of Mathematical Sciences, U.S. Air Force Academy, CO (SIAM)
John L. Scharf, Carroll College, Helena, MT
Lee Seitelman, Glastonbury, CT
Kathleen M. Shannon, Salisbury State University, Salisbury, MD (MAA)

Jonathan Shapiro, Dept. of Mathematics,
California Polytechnic State University, San Luis Obispo, CA
Robert M. Tardiff, Dept. of Mathematical Sciences,
Salisbury State University, Salisbury, MD
Michael Tortorella, Lucent Technologies, Holmdel, NJ
Marie Vanisko, Carroll College, Helena, MT (Triage)
Martin Wildberger, Electric Power Research Institute, Palo Alto, CA (SIAM)

Triage Judges
(all from Mathematics Dept., Carroll College, Helena, MT)
Head Triage Judge
Marie Vanisko
Associate Triage Judges
Mark Keefe, Terence J. Mullen, Phil Rose, and Jack Oberweiser

ICM

Contest Director
David C. Arney, Dept. of Mathematical Sciences, U.S. Military Academy

Elephant Population Problem

Head Judge
Gary W. Krahn, U.S. Military Academy, West Point, NY

Associate Judges
Kelly Black, Mathematics Dept., University of New Hampshire,
Durham, NH (Triage)
John Boland, Center for Industrial and Applied Mathematics (CIAM),
University of South Australia, Australia
Karen Bolinger, Dept. of Mathematics, Clarion University of Pennsylvania,
Clarion, PA
Ben Fusaro, Mathematics Dept., Florida State University, Tallahassee, FL (MAA)

Triage Judges
(all from Mathematics Dept., University of New Hampshire, Durham, NH)
Head Triage Judge
Kelly Black
Associate Judges
John B. Geddes, Gertrud Kraut, Dave Mecker, Jason Owen, Phil Ramsey, and Kevin Short

Sources of the Problems

Contributors of the problems were as follows:

- **Air Traffic Control Problem:** Robert Rovinsky, Federal Aviation Agency, Washington, DC
- **Channel Assignment Problem:** Jerrold R. Griggs, Dept. of Mathematics, University of South Carolina, Columbia, SC
- **Elephant Population Problem:** Anthony M. Starfield, Dept. of Ecology, Evolution, and Behavior, University of Minnesota, Minneapolis, MN

Acknowledgments

The MCM was funded this year by the National Security Agency, whose support we deeply appreciate. The ICM received major funding from the National Science Foundation. We thank Dr. Gene Berg of NSA for his coordinating efforts. The MCM is also indebted to INFORMS, SIAM, and the MAA, which provided judges and prizes.

I thank the MCM judges and MCM Board members for their valuable and unflagging efforts. Harvey Mudd College, its Mathematics Dept. staff, and Prof. Borrelli were gracious hosts to the judges.

Cautions

To the reader of research journals:

Usually a published paper has been presented to an audience, shown to colleagues, rewritten, checked by referees, revised, and edited by a journal editor. Each of the student papers here is the result of undergraduates working on a problem over a weekend; allowing substantial revision by the authors could give a false impression of accomplishment. So these papers are essentially au naturel. Light editing has taken place: minor errors have been corrected, wording has been altered for clarity or economy, and style has been adjusted to that of *The UMAP Journal*. Please peruse these student efforts in that context.

To the potential MCM Advisor:

It might be overpowering to encounter such output from a weekend of work by a small team of undergraduates, but these solution papers are highly atypical. A team that prepares and participates will have an enriching learning experience, independent of what any other team does.

Appendix: Successful Participants

KEY:

P = Successful Participation
H = Honorable Mention
M = Meritorious
O = Outstanding (published in this special issue)

A = Air Traffic Control Problem
B = Channel Assignment Problem
I = Elephant Population Problem

INSTITUTION	CITY	ADVISOR	A	B	I
ALABAMA					
Huntingdon College	Montgomery	Bob Robertson		P	
ALASKA					
Univ. of Alaska Fairbanks	Fairbanks	Chris Hartman	M	M	
CALIFORNIA					
Calif. Acad. of Math & Sci.	Carson	Brian R. Lawler			M
Calif. Lutheran University	Thousand Oaks	Cindy Wyels		P	
Calif. Poly. State Univ.	San Luis Obispo	Thomas O'Neil		O,H	
Calif. State U.	Bakersfield	Joseph R. Fiedler	P	P	
Calif. State U.	Northridge	Gholam-Ali Zakeri		P	
Calif. State U. Fullerton	Fullerton	Mario Martelli	P		H,P
Calif. State U. Monterey Bay	Seaside	Dan Fernandez and Michael Dalton	C	P	
Harvey Mudd College	Claremont	Michael Moody	H	M	O,P
		Zachary Dodds	M	H	
Humboldt State Univ.	Arcata	Roland Lamberson	P		
Sonoma State University	Rohnert Park	Sunil K. Tiwari	P		
Univ. of Calif. - Berkeley	Berkeley	Rainer K. Sachs	H	P	
COLORADO					
Colorado College	Colorado Springs	Jane McDougall		H	
		Jennifer Courter		H	
Mesa State College	Grand Junction	Bill Tiernan		P,P	
		Edward Bonan-Hamada	H,H		
Regis University	Denver	Linda Duchrow	P	P	
U.S. Air Force Academy	USAF Academy	Dawn Stewart		P,P	
Univ. of Colorado	Colorado Springs	Jon Epperson	M	H	
	Boulder	Anne M. Dougherty	O		
		Anne Dougherty and Bengt Fornberg			H
Univ. of Southern Colorado	Pueblo	James Louisell		P	
CONNECTICUT					
Sacred Heart Univ.	Fairfield	Antonio A. Magliaro	P		
Southern Conn. State Univ.	New Haven	Ross B. Gingrich		P	
		Theresa Bennett		P	

INSTITUTION	CITY	ADVISOR	A	B	I
Univ. of Bridgeport	Bridgeport	Dr. Natalia Romalis	P		
U.S. Coast Guard Academy	New London	John Freda			P
Western Conn. State Univ.	Danbury	C Edward Sandifer	H,P		
Yale University	New Haven	Steven Orszag		M	
DISTRICT OF COLUMBIA					
Georgetown University	Washington	Andrew Vogt	P	P	
FLORIDA					
Florida A&M University	Tallahassee	Bruno Guerrieri		P	
Florida Inst. of Tech.	Melbourne	Gary W. Howell		P	
Jacksonville Univ.	Jacksonville	Robert A. Hollister	H	P	
Stetson University	Deland	Lisa O. Coulter	M,H		
Univ. of North Florida	Jacksonville	Peter A. Braza		P	
GEORGIA					
Agnes Scott College	Decatur	Robert A. Leslie		P	
Georgia Southern Univ.	Statesboro	Eric Funasaki		P	
		Dr. Goran Lesaja	P		
		Gary Huband	H		
State Univ. of West Georgia	Carrollton	Scott Gordon	H		
HAWAII					
Kapi'olani Community	Honolulu	Susan Moore	P		
IDAHO					
Boise State University	Boise	Stephen H. Brill			P
ILLINOIS					
Greenville College	Greenville	Galen R. Peters	P	P	
Northern Illinois University	Dekalb	Hamid Bellout		P	
Wheaton College	Wheaton	Paul Isihara		P,P	
INDIANA					
Goshen College	Goshen	David Housman	P	P	
		David Housman and Patricia Oakley			P
Indiana University	Bloomington	John Brothers		P,P	
Rose-Hulman Inst. of Tech.	Terre Haute	Dr. Frank Young	M		
		Frank Young		P	
		David Rader		M,P	
		Sharon A. Jones and Robert J. Houghtalen			H,P
Saint Mary's College	Notre Dame	Peter D. Smith		H,P	

INSTITUTION	CITY	ADVISOR	A	B	I
IOWA					
Drake University	Des Moines	Alexander F. Kleiner	M	P	
Graceland College	Lamoni	Ronald K. Smith	H		
		Steve K. Murdock	P	P	
Grand View College	Des Moines	Timothy L. Hardy	P		
Grinnell College	Grinnell	Marc Chamberland		M,H	
		Mark Montgomery		P,P	
Iowa State University	Ames	Stephen J. Willson		H	
Luther College	Decorah	Reginald D. Laursen		M,H	
Mt. Mercy College	Cedar Rapids	Kent R. Knopp		M,P	
Simpson College	Indianola	M.E. Waggoner	M,P		
		Randy Bower	P	P	
Univ. of Northern Iowa	Cedar Falls	Gregory M. Dotseth	P		
Wartburg College	Waverly	Mariah Birgen	P		
KANSAS					
Benedictine College	Atchison	Jo Ann Fellin, OSB		P	
KENTUCKY					
Asbury College	Wilmore	Kenneth P. Rietz		M	
Bellarmine College	Louisville	Bill Hardin	H		
LOUISIANA					
Northwestern State U. of La.	Natchitoches	Richard DeVault	P		
MAINE					
Bowdoin College	Brunswick	Adam B. Levy		P	
Colby College	Waterville	Jan Holly	P	H	
MARYLAND					
Goucher College	Baltimore	Robert E. Lewand		P	
Mt. St. Mary's College	Emmitsburg	Bill O'Toole	P		
Salisbury State Univ.	Salisbury	Michael Bardzell		H	
		Steven M. Hetzler	H		
MASSACHUSETTS					
MIT	Cambridge	Michael P. Brenner		M	
		Michael P. Brenner and L. Mahadevian			M,H
Simon's Rock College	Great Barrington	Allen B. Altman	P	P	
		Michael Bergman		P	
Smith College	Northampton	Ruth Haas	P	P	
Univ. of Massachusetts	Lowell	James K. Graham-Eagle		P	
		Lou Rossi	P		
	Amherst	Robert B. Kusner		P	

INSTITUTION	CITY	ADVISOR	A	B	I
Western New England College	Springfield	Lorna Hanes		P	
Worcester Poly. Inst.	Worcester	Bogdan Vernescu	M		
		Richard Jordan		H	
MICHIGAN					
Calvin College	Grand Rapids	Dorothea Pronk		M	
Eastern Michigan University	Ypsilanti	Christopher E. Hee		H	
Hillsdale College	Hillsdale	John P. Boardman	P	H	
Lawrence Tech. Univ.	Southfield	Howard Whitston	P		
		Ruth G. Favro		P	
		Scott Schneider		H	
Michigan State Univ.	E. Lansing	Charles R. MacCluer		H	
Univ. of Michigan	Dearborn	David James		M	
MINNESOTA					
Macalester College	St. Paul	Daniel A. Schwalbe	P		
		Daniel Kaplan	P	P	
MISSOURI					
Crowder Coll.	Neosho	Cheryl Ingram			P
		Patrick Cassens	P		
Northwest Missouri State Univ.	Maryville	Russell Euler	P		
St. Louis University	St. Louis	David A. Jackson	H		
Truman State University	Kirksville	Steve Smith		P,P	
Washington University	St. Louis	Hiro Mukai		O	
Wentworth Military Academy	Lexington	Jacque Maxwell	H		
MONTANTA					
Carroll College	Helena	Jack Oberweiser	P		
		Mary Keeffe	P		
		Phil Rose	P		
		Terry Mullen			P
NEBRASKA					
Hastings College	Hastings	David B. Cooke	P,P		
NEVADA					
Sierra Nevada College	Incline Village	Steve Ellsworth			P
Universuty of Nevada	Reno	Mark M. Meerschaert	P		
NEW JERSEY					
Rowan University	Glassboro	Hieu Nguyen		P	
		Paul J. Laumakis	P		
		Sam Lofland			P

INSTITUTION	CITY	ADVISOR	A	B	I
NEW MEXICO					
New Mexico Inst. of Mining & Tech.	Socorro	William D. Stone	P		
New Mexico State University	Las Cruces	Marcus S. Cohen	P		
NEW YORK					
Colgate University	Hamilton	Thomas W. Tucker		H,P	
Ithaca College	Ithaca	James E. Conklin		P	
	John C. Maceli		H		
Nazareth College	Rochester	Kelly M. Fuller	H		
Rensselaer Poly. Inst.	Troy	Bruce Piper	P		
		Donald A. Drew		P	
Roberts Wesleyan College	Rochester	Carlos A. Pereira			P
St. Bonaventure Univ.	St. Bonaventure	Albert G. White	P		
		Maureen Cox		P,P	
SUNY Cortland	Cortland	R. Bruce Mattingly and R. Lawrence Klotz			P
U.S. Military Academy	West Point	David Bailey	O		
		Diane Nelson	H		
		Greg Parnell		M	
		Michael Jaye and Greg Fleming			M
		Michael Phillips and Michael Darrow			P
		James S. Rolf		H	
Wells College	Aurora	Carol C. Shilepsky		P	
Westchester Comm. College	Valhalla	Sheela Whelan	P	P	
NORTH CAROLINA					
Brevard College	Brevard	Theresa A. Bright	P		
Duke University	Durham	David P. Kraines	O		
Elon College	Elon College	Todd Lee			P
N.C. School of Sci. & Math.	Durham	Dot Doyle and Dan Teague			O,O
North Carolina State Univ.	Raleigh	Ranji S. Ranjithan			H
Univ. of North Carolina	Chapel Hill	Jon W. Tolle	P		
	Wilmington	Russell L. Herman	P	H	
Wake Forest University	Winston-Salem	Edward Allen		O	
Western Carolina Univ.	Cullowhee	Jeff Graham		H	
OHIO					
Capital University	Columbus	Dr. Ignatios Vakalis	P		
Hiram College	Hiram	Angela Spalsbury		H,P	
Marietta College	Marietta	Tom LaFramboise	P		
Miami University	Oxford	Douglas E. Ward		P	
Ohio University	Athens	David N. Keck			H

INSTITUTION	CITY	ADVISOR	A	B	I
The College of Wooster	Wooster	Pamela Pierce	H		
University of Cincinnati	Cincinnati	Charles Groetsch	M		
Xavier University	Cincinnati	Richard J. Pulskamp	P,P		
Youngstown State Univ.	Youngstown	Bob Kramer		P	
		Scott Martin			M
		Steve Hanzely	M		
		Thomas Smotter	M	P	P
OKLAHOMA					
Univ. of Central Oklahoma	Edmond	Charles L. Cooper	H		
		Daniel J. Endres		P	
		Charlotte Simmons and Jesse Byrne			P
OREGON					
Eastern Oregon State College	LaGrande	David Allen	P		
		John Thurber	P		
		Anthony Tovar and Jeffrey Putnam			P
		Richard Hermens and Tom Herrmann			P
Lewis & Clark College	Portland	Robert W. Owens		O,P	
Southern Oregon State College	Ashland	Kemble R. Yates		P	
PENNSYLVANIA					
Bloomsburg University	Bloomsburg	Kevin Ferland	P		
		Kevin Ferland and Scott Inch			M
Bucknell University	Lewisburg	Sally Koutsoliotas	P		
Chatham College	Pittsburgh	Eric Rawdon		P	
		Larry Viehland	P		
Clarion University	Clarion	Andrew Turner and Sharon Challener			P
		John W. Heard		H	
		Jon Beal		P	
		William D. Krugh			P
Gettysburg College	Gettysburg	James P. Fink	P	M	
		John Jaroma	H		
		Sharon Stephenson		H	
Haverford College	Haverford	Robert Manning		P	
Lafayette College	Easton	Thomas Hill	M		
Messiah College	Grantham	Douglas C. Phillippy		P	
		James Makowski		H	

INSTITUTION	CITY	ADVISOR	A	B	I
RHODE ISLAND					
Rhode Island College	Providence	David L. Abrahamson	P		
SOUTH CAROLINA					
Central Carolina Tech. Coll.	Sumter	Michael L. Salais	P		
Charleston Southern Univ.	Charleston	Stan Perrine		P,P	
Univ. of South Carolina	Aiken	Laurene Fausett	P		
SOUTH DAKOTA					
Northern State Univ.	Aberdeen	A.S. Elkhader	H		
TENNESSEE					
Lipscomb University	Nashville	Gary Hall	H		
		Mark Miller	P		
TEXAS					
Abilene Christian University	Abilene	David Hendricks		P	
Angelo State Univ.	San Angelo	John C. (Trey) Smith		H	
Austin Community College/River	Austin	Allison Sutton		P	
Southwestern Univ.	Georgetown	Therese Shelton		P	
Stephen F. Austin State Univ.	Nacogdoches	Colin Starr	P		
Trinity University	San Antonio	Allen Holder		M	
		Jeffrey Oldham		P,P	
		Tarynn Witten	M		
University of Dallas	Irving	Pete McGill		P	
University of North Texas	Denton	John Quintanilla	H		
University of Texas at Dallas	Richardson	Tiberiu Constantinescu	P		
UTAH					
University of Utah	Salt Lake City	Don H. Tucker		H	
Weber State University	Ogden	Richard Miller	H		
VERMONT					
Johnson State College	Johnson	Glenn Sproul	P,P		
VIRGINIA					
Governor's School	Richmond	Crista Hamilton	O	H	
		John Barnes		H	
James Madison University	Harrisonburg	Caroline Smith	H		
		James S. Sochacki		P	
Randolph-Macon College	Ashland	Eve Torrence		P	
University of Richmond	Richmond	Kathy W. Hoke		M	
Virginia Western Comm. College	Roanoke	Ruth Sherman		P	

INSTITUTION	CITY	ADVISOR	A	B	I
WASHINGTON					
Pacific Lutheran University	Tacoma	Rachid Benkhalti		M,P	
University of Puget Sound	Tacoma	Robert A. Beezer		H,P	
University of Washington	Seattle	Peter Schmid		P	
		Randall J. LeVeque		P	
Western Washington University	Bellingham	Igor Averbakh	H		
WISCONSIN					
Beloit College	Beloit	Paul J. Campbell	H		P
Cardinal Stritch University	Milwaukee	Sr. Barbara Reynolds		P	
Ripon College	Ripon	David Scott	P	P	
St. Norbert College	De Pere	John A. Frohliger		P	
Univ. of Wisconsin	Platteville	Sherrie Nicol		P	
	Stevens Point	Nathan Wetzel		P	
Wisconsin Lutheran College	Milwaukee	Marvin C. Papenfuss		P	
AUSTRALIA					
University of New South Wales	Sydney	James Franklin		M,H	
		Bruce Henry	H,P		
University of Southern Queensland	Toowoomba	Tony Roberts		P	
CANADA					
Brandon University	Brandon, MB	Doug Pickering	P	P	H
Memorial Univ. of Newfoundland	St. John's, NFLD	Andy Foster	P	H	
Okanagan University College	Kelowna BC	Dr. Heinz Bauschke	H		
University of Ottawa	Ottawa, ON	Luc Demers	H,P		
University of Saskatchewan	Saskatoon, SK	James A. Brooke		P	
		Raj Srinivasan	M		
		Tom Steele		P	
University of Toronto	Toronto ON	Nicholas A. Derzko		M,M	P
York University	Toronto ON	Juris Steprans	P		
		Neal Madras		H	
University of Western Ontario	London, ON	Peter H. Poole	P	P	
CHINA					
Anhui University	Hefei, Anhui	Chen Junsheng		M	
		Wang Dapeng		H	
		Wang HaiXi'an		H	
Beijing Institute of Technology	Beijing	Xiao Dicui	H	P	
Beijing Union University	Beijing	Ren Kailong		P	
		Xing Chunfeng		H	
		Zeng Qingli	H		
Beijing Univ. of Chemical Tech.	Beijing	Liu Damin		H	
		Zhao Baoyuan		H	

INSTITUTION	CITY	ADVISOR	A	B	I
Beijing U. of Post & Telecomm.	Beijing	He Zuguo		M	
		Sun Hongxiang		M,P	
Central-South Univ. of Tech.	Changsha, Hunan	Zheng Zhoushun		P	
		Zhang Hongyan		H,P	
China U. of Mining & Tech.	Xuzhou, Jiangsu	Zhou Shengwu			M
		Wu Zongxiang		M	
		Xue Xiuqian		H	
		Zhu Kaiyong		H	
Chongqing University	Chongqing	Chen Yihua	M		
		Gong Qu		H	
		He Renbin and Zhang Dan			H
		Li Chuandong		P	
		Liu Qiongsun			H
		Zhao Pingyong		H	
Coll. of Sciences, Jiamusi Univ.	Jiamusi, Heilongjiang	Shan Baifeng		P,P	
		Gu Lizhi	P	P	
Dept. of System Science	Changsha, Hunan	Duan Xiaolong		P	
East China U. of Sci. & Tech.	Shanghai	Shi Jinsong		M	
		Shao Nianci	H		
		Lu Xiwen		M	
		Lu Yuanhong	M		
		Lu Xiwen and Shao Nianci			P
		Lu Yuanhong and Dong Liming			H
First Middle School of Jiading	Shanghai	Chen Gan		H	
		Fang Yunping		M	
Fudan University	Shanghai	Cao Yuan and Cai Zhijie			H
		Xu Qinfeng		P	
		Gong Xueqing		P,P	
		Cai Zhijie		M	
		Cai Zhijie and Cao Yuan			P
Guangdong Commercial Coll.	Guangzhou, Guangdong	Chen Guanghui		P	
Harbin Engineering Univ.	Harbin, Heilongjiang	Luo Yuesheng		H	
		Shen Jihong		P	
		Zhang Xiaowei		H	
		Shang Shouting and	H	P	
		Shang Shouting Zheng Tong			H

INSTITUTION	CITY	ADVISOR	A	B	I
Harbin Institute of Tech.	Harbin, Heilongjiang	Shang Shouting and Yu Xiujuan			P
		Wang Xuefeng		M	
		Zhang Chiping	H		
Hefei University of Tech.	Hefei, Anhui	Bao Jie		P	
		Du Xueqiao		P	
		Huang Youdu	P		
		Su Huaming			H
		Yang Youqing			P
		Zhou Yongwu		P	
Huazhong University	Wuhan, Hubei	Wang Yizhi		P,P	
Inst. of Info. Sci. & Eng.	Shenyang, Liaoning	Cui Jianjiang		H	
		Hao Peifeng		H	
		Xiao Wendong		M	
Inst. of Operations Research	Qufu	Shi Zhenjun		P	
Jilin University	Changchun, Jilin	Huang Qingdao		P	
		Lu Xian Rui		P	
Jinan University	Guangzhou, Guangdong	Fan Suohai		P	
		Ye Shiqi	P		
Nanjing Normal University	Nanjing, Jiangsu	Fu Shitai		P,P	
		Yao Tianxing		H,H	
Nankai University	Tianjin	Ruan Jishou		H	P
		Chen Qiushuang		H	
		Xingwei Zhou		P	
		Chen Zengqiang	H		
National U. of Defence Tech.	Changsha, Hunan	Cheng Lizhi		M,H	
		Wu Mengda		O	
Northwest Textile Inst.	Xi'an, Shaanxi	He Xingshi		P	
Northwest University	Xi'an, Shaanxi	He Ruichan		H	
Northwestern Polytech. U.	Xi'an, Shaanxi	Liu Xiaodong		M	
		Liu Xiaodong and Peng Guohua			P
		Peng Guohua		M	
		Wang Mingyu	H		
		Xu Wei	H		
		Xu Wei and Wang Mingyu			M
Peking University	Beijing	Deng Minghua	M	P	
		Lei Gongyan		M,H	
		Shao Min and Zhang Tao			M,H
Science School of Xi'an Jiaotong U.	Xi'an, Shaanxi	Zhou Yicang			H
		He Xiaoliang	M		
		He Xiaoliang	P		

INSTITUTION	CITY	ADVISOR	A	B	I
Shandong University	Jinan, Shandong	Wang Dongsheng		P	
Shanghai Jiao Tong Univ.	Shanghai	Gong Peimin and Liu Xiaojun			P,P
		Huang Jianguo		H	
		Lei Zhishu	H		
		Song Baorui		P	
		Zhou Gang		M	
Shanghai Foreign Languages School	Shanghai	Pan Lioun		P	
		Wan Baihe		M	
Shanghai Normal Univ.	Shanghai	Zhu Detong		P,P	
		Guo Shenghuan		P	
Sichuan University	Chengdu, Sichuan	Zou Shuchao	P		
		Zhou Jie		P	
		Yang Zhihuo	P		
South China Univ. of Tech.	Guangzhou, Guangdong	He Chunxiong and He Dehua			H
		Hao Zhifeng		P	
		Xu Shuwen and Hong Yi			P
		Fu Hongzhuo		M	
		Xie Lejun	H		
		He Chunxiong	P		
Southeast University	Nanjing	Chen Enshui			M
		Huang Jun	H	M,M,P	
Tianjin University	Tianjin	Rong Ximin		P	
		Liu Zeyi	H		
		Hu Zhiming		M,H	
		Hu Zhiming and Mei Helu			P
Tsinghua University	Beijing	Ye Jun	H	P	
Univ. of Elec. Sci. & Tech.	Chengdu	Wang Jianguo		P	
		Xu Quanzhi	H	H,P	
		Zhong Erjie	H		
Univ. of Sci. & Tech. of China	Hefei, Anhui	Wan Qian			M
		Xu Jizheng			H
		Hong Quan		H	
		Jin Letian		P	
		Sun Liang	M		
		Wei Liu		H	
Wuhan U. of Hydraulics & Eng.	Wuhan, Hubei	Peng Zuzeng		M	
		Cheng Guixing		H	
		Huang Chongchao		H	
		Li Yuhong		H	

INSTITUTION	CITY	ADVISOR	A	B	I
Xi'an Jiaotong University	Xi'an, Shaanxi	Dai Yonghong			P
		Zhou Yicang		P,P	
Xi'an Univ. of Tech.	Xi'an, Shaanxi	Du Zhanrong		P	
		Tang Ping	P		
Xidian School of Computers	Xi'an, Shaanxi	Zhou Shuisheng and Liu Hongying			P
Xidian University	Xi'an, Shaanxi	Liu Hongwei		H	
		Zhang Zhuokui		H	
		Mao Yongcai		P	
Zhejiang University	Hangzhou, Zhejiang	Yang Qifan	M	M	
		He Yong	M	M	
		Cong Zhang, Chu Jiaowu, and He Yong			M
		Cong Zhang, Chu Jiaowu, and Yang Qifan			M
Zhong Shan University	Guangzhou, Guangdong	Chen Zepeng		H	
		Yin Zhaoyang		P	
		Li Caiwei		P	
		Li Chaorui		H	
		Tang Mengxi			H
		Yuan Zhuojian			H
FINLAND					
Paivola College	Tarttila	Esa Lappi	P	M	
HONG KONG					
Hong Kong Baptist Univ.	Kowloon Tong	Chong Sze Tong		H	
		Wai Chees Shiu		P	
IRELAND					
University College Cork	Cork	James J. Grannell	P		
		Michael Quinlan	M		
		Patrick Fitzpatrick	P		
		Martin Stynes		H	
University College Dublin	Belfield, Dublin	E.A. Cox		P	
		Peter Duffy		H	
		Fergus Gaines	H		
		Niamh O'Sullivan		P	
University College Galway	Galway	Martin Meere	H		
		Michael P. Tuite		H	
University of Limerick	Limerick	Gordon S. Lessells		P	

INSTITUTION	CITY	ADVISOR	A	B	I
LITHUANIA					
Vilnius University	Vilnius	Algirdas Zabulionis		P	
		Antanas Mitasiunas		P	
		Ricardas Kudzma		P	
		Tadas Meskauskas	P		
SOUTH AFRICA					
University of Stellenbosch	Matieland	Jan H. van Vuuren	P	P	

Acknowledgment

The editor thanks Chia Tzun Goh ’01 of Beloit College for his help in identifying family names of team advisors from China, for whom the family name is listed first.

Air Traffic Control

Samuel Westmoreland Malone
Jeffrey Abraham Mermin
Daniel Bertrand Neill
Duke University
Durham, NC

Advisor: David P. Kraines

Introduction

We propose five models. Each gives a metric for measuring *eventual danger* and a threshold such that controller intervention becomes necessary. An *immediate danger* metric prioritizes problems for a controller.

We test the models on several different sample cases. The Close-Approach model matches most closely our intuitive understanding of the situations.

We present an algorithm that models the decision process of a controller detecting and solving conflicts. The time-complexity of scanning for potential conflicts varies quadratically with the number of airplanes, though this could be reduced by clustering the airplanes by proximity. We argue that conflict resolution for a cluster of n airplanes is an NP-complete problem with worst-case complexity $2^{n(n+1)/2}$.

Assumptions and Hypotheses

- A *near mid-air collision* (sometimes called a *near miss*) is defined by the FAA as an incident in which two airplanes pass within 500 ft of each other [Federal Aviation Administration 2000].
- The airspace can be represented by a convex subset of R^3.
- Air-traffic controllers have established protocols to prevent airplanes from colliding when crossing airspace boundaries in opposite directions.
- We know the position and velocity of every airplane in the airspace, with negligible error.

The UMAP Journal 21 (3) (2000) 241–256.

- Every airplane has sufficiently negligible acceleration that linear models for its movement make sense over at least the next 2 min unless it is:
  - accelerating under the direction of a controller (so the controller has already determined potential conflicts),
  - taking off (so it is not in the airspace used by cruising airplanes), or
  - attempting to land (so it is not in the airspace used by cruising airplanes).
- Airplanes can accelerate through a 2-minute turn, either clockwise or counterclockwise, parallel to the xy-plane [Denker 2000].
- Airplanes travel within vertically well-separated planes, called "cruising altitudes" [Mahalingam 1999, 27]. Airplanes in different cruising altitudes pose no danger to each other.
- Airplanes can accelerate at a given maximum rate in the direction of their velocity. (In particular, they do not cruise at their maximum speed.)
- Two airplanes present an *eventual danger* when, if their velocities are allowed to go unchanged indefinitely,
  - they will collide,
  - they will pass "near" each other at some time, or
  - they will pass through "nearby" points in space at "nearby" times.

  We later discuss appropriate values of "near."

Possible Solutions to the Danger Problem

Two considerations dominate in determining the danger of two airplanes to each other: the proximity that they will attain and the time until it occurs. We present several approaches:

- The Trivial Model provides yes-or-no answers to the question "Will the airplanes collide?"
- The Probabilistic Model determines risk based on the probabilities of collisions and near misses.
- The Close-Approach Model calculates the closest approach of two airplanes and the time until it occurs.
- The Space-Time Model considers the closest approach in four-dimensional space-time and the time until it occurs.
- The Logarithmic Derivative Model approximates a human observer's intuitions about how fast the airplanes are approaching each other.

The Trivial Model

How It Works

The Trivial Model ignores effects such as wind, measurement uncertainty, or piloting imperfection, which could make an airplane's course deviate from the linear projection of its current position and velocity.

Large commercial aircraft have lengths and wingspans of about 200 feet. Thus, *a collision occurs only if the centers of airplanes pass within 200 ft of each another, and a near miss occurs only if they pass within 700 ft*.

Suppose airplanes A and B have position and velocity vectors p_A, v_A and p_B, v_B. Set $p = p_B - p_A$ and $v = v_B - v_A$, the position and velocity of B relative to A. The distance of closest approach is the altitude from A to v (**Figure 1**). Its length is $d = |p| \sin\theta = |p \times v| \,/\, |v|$.

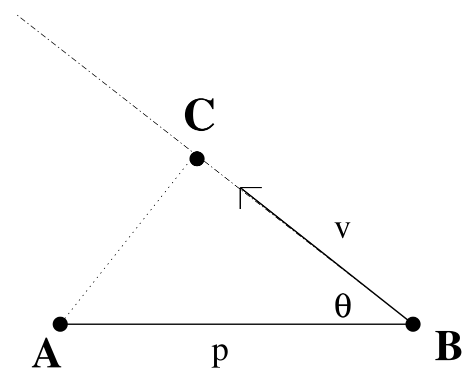


Figure 1. The position and velocity vectors of airplane B relative to airplane A.

There will be a collision if d is ever less than 200 ft, and a near miss if it is ever less than 700 ft. A measure of the eventual danger takes on three discrete values a ($\gg 1$), 1, or 0 corresponding to a collision, a near miss, and no danger. The value of a is best determined empirically.

Strengths and Weaknesses

This model is simple and efficient. However, it assumes that airplanes always travel at a constant speed in a straight line; but in fact airplanes are buffeted by changing winds and their actual trajectories vary significantly and chaotically from those predicted by a linear model. Additionally, this model considers only eventual danger, not how soon immediate danger will be present, though the model could be extended to rank collisions and near misses based on immediate danger (time to collision or near miss).

Probabilistic Simulation Model

How It Works

The Probabilistic Simulation Model calculates the probability that a given situation will result in a collision or near miss by using a Monte Carlo simulation. To do so, it performs a large number of random trials (each of which may result in a collision, near miss, or neither) and computes a measure of eventual danger: danger = $c_1 x + c_2 y$, where x and y are the probabilities of collision and near miss and, as in the Trivial Model, we set $c_2 = 1$ and $c_1 \gg 1$.

We assume normal distributions of each airplane's speed and direction, with the mean the measured value and the standard deviation specified by the user. For each trial, a normally distributed random value of each quantity is chosen, then both airplanes' paths are extrapolated linearly; we use the minimum-distance formula from the Trivial Model to determine whether a collision, near miss, or neither occurs.

Strengths and Weaknesses

Though a minimum time to collision or near miss is computed, this time is not taken into account in the measure of danger. Hence, we add an optional user-specified maximum time horizon; if two airplanes do not reach their minimum distance by then, their distance at that time is considered instead of the (later) minimum distance. Thus, we ignore conflicts that occur far in the future, focusing on more immediate dangers. According to one source [MAICA: MET improvement . . . , 124], conflict analysis tools that extrapolate based on current aircraft trajectories "operate over a short time horizon, generally less than two minutes."

The Close-Approach Model

How It Works

We expect eventual danger to be inversely related to closest approach and immediate danger to be inversely related to time until that closest approach. Thus, we have, as a first approximation,

$$\text{Eventual danger} \approx \frac{1}{(\text{distance of closest approach})^{\alpha}},$$

$$\text{Immediate danger} \approx \frac{1}{(\text{distance of closest approach})^{\alpha}\,(\text{time until closest approach})^{\beta}}.$$

Since danger could be averted by accelerating the airplanes away from each other, the extra separation achieved should be proportional to the square of the time of acceleration. Since the time during which they can accelerate is bounded by the time until projected close approach, it seems reasonable to set $\beta = 2\alpha$. Since raising to a positive power doesn't affect ordering, we set $\beta = 2$, $\alpha = 1$.

Such a simple formula runs into trouble in boundary situations:

- No matter how far away the airplanes will be at their closest approach, immediate danger goes to infinity as they come to closest approach.
- If the two airplanes are on a collision course, the formula gives infinite immediate danger, no matter how much time remains until collision.
- If the airplanes have nearly equal velocities, the formula rates immediate danger as near zero (unless the aircraft are practically on top of each other) when it should intuitively be inversely proportional to current separation.

We fix the formula as follows:

$$\text{Immediate danger} = \frac{1}{(\text{distance of closest approach} + c_1)\,(\text{time until closest approach} + c_2)^2} + \frac{c_3}{\text{current separation}},$$

where c_1, c_2 and c_3 are positive constants, probably best determined empirically.

Now we calculate the ingredients in the formula. If the distance of closest approach of a pair of airplanes is sufficiently large (e.g., $d > 5$ nautical mi [Mahalingam 1999, 26–27]), they pose no danger to each other. If they will pass closer, we rate the level of eventual danger as d.

The time until closest approach is the time until airplane B reaches point C:

$$\frac{\left|\overrightarrow{BC}\right|}{|v|} = \frac{|p|\cos\theta}{|v|} = \frac{\dfrac{p \cdot v}{|v|}}{|v|} = \frac{p \cdot v}{|v|^2}.$$

Plugging in, we get

$$\text{Immediate danger} = \frac{|v|^5}{(|p \times v| + |v|\,c_1)\left(p \cdot v + |v|^2\,c_2\right)^2} + \frac{c_3}{|p|}.$$

Since we get $0/0$ in the first summand when $v = 0$, that is, when the two airplanes are flying parallel to each other, in this case we set the immediate danger equal to $c_3/|p|$.

Strengths and Weaknesses

The danger can be computed with just over 50 basic numeric operations if there is eventual danger and about half as many to conclude that there is not; thus, a personal computer could handle this computation every second for several thousand airplane-pairs, or about 500 airplanes every 2 min.

This model does not worry about airplanes that pass near each other in time but not in space. For example, it cannot distinguish between two airplanes flying the exact same route through space with a time-separation of 15 sec and two airplanes flying parallel with a physical separation of 2 mi in a direction orthogonal to their velocity; the first situation appears to us to be much more dangerous. The next model attempts to differentiate such situations.

The Space-Time Model

How It Works

The Space-Time Model uses similar reasoning to the Close-Approach Model but considers the airplanes' proximity in space-time rather than simply in physical space. Thus, intuitively, we have:

$$\text{Eventual danger} \approx \frac{1}{\text{closest approach in space-time}},$$

$$\text{Immediate danger} \approx \frac{1}{\left(\text{closest approach in space-time}\right)\left(\text{time until closest approach}\right)^2}.$$

The same corrections for boundary conditions apply, leaving

$$\text{Immediate danger} = \frac{1}{\left(\text{closest approach in space-time} + \gamma_1\right)\left(\text{time until closest approach} + \gamma_2\right)^2} + \frac{\gamma_3}{\text{current space-time separation}}.$$

These quantities are harder to compute than in the Close-Approach Model. We can represent the future of airplane A (initially at the origin) and airplane B by rays in R^4, parametrized by the vectors $\left(v_{a_x}t_a, v_{a_y}t_a, v_{a_z}t_a, kt_a\right)$ and $\left(v_{b_x}t_b + p_x, v_{b_y}t_b + p_y, v_{b_z}t_b + p_z, kt_b\right)$, for $t_a, t_b > 0$, where k is a constant chosen so that one unit of time is as dangerous as one unit in one of distance. Mahalingam [1999, 26–27] equates a 15-min separation to a 5-nautical-mile separation; we assume that this equivalence scales. Then k equals 5 nautical mi per 15 min, or about 34 ft/s.

For any t_a, t_b, the space-time distance between airplane A at time t_a and the airplane B at time t_b is

$$\delta(t_a, t_b) = \left|\left(v_{a_x} t_a, v_{a_y} t_a, v_{a_z} t_a, k t_a\right) - \left(v_{b_x} t_b + p_x, v_{b_y} t_b + p_y, v_{b_z} t_b + p_z, k t_b\right)\right|.$$

This yields

$$\left(\delta(t_a, t_b)\right)^2 = A t_a^2 + B t_a t_b + C t_b^2 + D t_a + E t_b + |p|^2,$$

where

$$A = k^2 + |v_a|^2, \; B = -2k^2 - 2v_a \cdot v_b, \; C = k^2 + |v_b|^2, \; D = -2p \cdot v_a, \; E = 2p \cdot v_b.$$

The minimum space-time distance occurs at t_a and t_b that minimize this expression, that is, where $\nabla \delta^2 = 0$:

$$t_\alpha = \frac{2CD - BE}{B^2 - 4AC}, \qquad t_\beta = \frac{2AE - BD}{B^2 - 4AC},$$

which are well-defined whenever the velocities are not equal, since

$$B^2 - 4AC = 4\left[(v_a \cdot v_b)^2 - |v_a|^2 |v_b|^2 + 2k^2 v_a \cdot v_b - k^2 |v_a|^2 - k^2 |v_b|^2\right].$$

The first two terms add to less than zero by Cauchy-Schwarz, and the last three by AM-GM, with equality in both cases only when the velocities are equal. [The case when the velocities are equal is handled more simply: For every t_α, there is a unique t_β satisfying $B t_\alpha + 2C t_\beta + E = 0$ (since C is always positive) that yields the minimum space-time separation.] The minimum space-time separation is $\delta(t_\alpha, t_\beta)$, and the time until this separation is $\min\{t_\alpha, t_\beta\}$.

The current space-time separation would appear to be the minimum of the space-time separations between the current position of airplane A and the future of airplane B, and the current position of B and the future of A.

Determination of danger is done as in the Close-Approach Model, except that the times associated to closest approach must be computed first. If either is negative, then any danger posed by this airplane-pair has already been avoided, and so the eventual and immediate dangers are set to zero. Then the space-time separation of the closest approach is computed; if it is sufficiently large (e.g., more than 5 nautical mi), then the airplanes pose no danger to each other.

Strengths and Weaknesses

Every airplane-pair receives at least as high immediate- and eventual-danger measures from the Space-Time Model as from the Close-Approach Model, while the Space-Time Model recognizes as dangerous some cases that the Close-Approach Model does not. The Space-Time Model is not significantly slower in operation.

On the other hand, this model is much more opaque to any human who must try to work with it; human beings are not equipped to think in terms of extra dimensions.

The Logarithmic Derivative Model

How It Works

The model arises from the observation that, if the velocities of airplanes A and B remain constant, the time derivative dd/dt of the distance between the airplanes is monotonically increasing with time (unless the airplanes are traveling in the same line, in which case it is constant) and is bounded both above and below. Thus, the negative derivative is monotonically decreasing, and

$$-\left.\frac{d-\ell}{\frac{dd}{dt}}\right|_{t=t_0}$$

gives a lower bound on the time between t_0 and any future time t at which the airplanes are separated by a distance less than some danger threshold ℓ. Thus, the reciprocal of this quantity,

$$-\frac{\frac{dd}{dt}}{d-\ell}$$

(the negative derivative of the logarithm of $d-\ell$), might work as a measure of immediate danger.

We investigate the behavior of this function. Suppose that airplane B has position and velocity vectors p and v in some frame of reference where A is stationary at the origin, as in **Figure 2**.

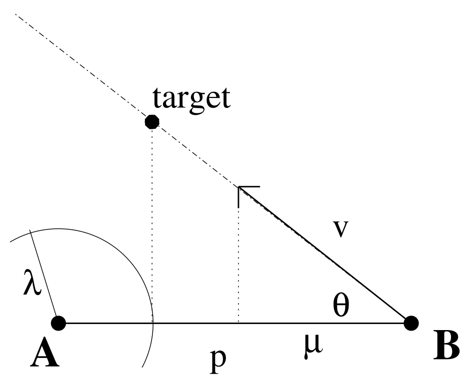


Figure 2. The logarithmic derivative.

Now

$$\frac{d}{dt}\left(|p|^2\right)=\lim_{h\to 0}\frac{|p+hv|^2-|p|^2}{h}=\lim_{h\to 0}\frac{2hp\cdot v+h^2|v|^2}{h}=2p\cdot v.$$

But

$$\frac{d}{dt}\left(|p|^2\right)=2|p|\frac{d}{dt}|p|,$$

so

$$\frac{d}{dt}|p|=\frac{p\cdot v}{|p|}=|v|\cos\theta.$$

This quantity is represented in **Figure 2** by μ, the projection of one time-unit of velocity onto p. Dividing by $|p| - \ell$ gives the number of time units until the projection onto p intersects the circle of radius ℓ about A.

Consider also the behavior of this function as time passes (**Figure 3**):

- If B will not pass within ℓ of A, the target point goes to infinity as B approaches the point of least separation from A; the measure simultaneously goes to zero, then becomes negative as that point is crossed. This is fine; all danger is past once the point of least separation is reached.
- If B will pass through the circle of radius ℓ, the target point converges to the intersection of B's trajectory with the circle; the measure goes to infinity as B approaches the circle, then becomes negative once it passes inside. If ℓ is sufficiently small (e.g., 700 ft, the threshold for a near miss), this is also fine: Once the two airplanes are within ℓ of each other, it is already too late.

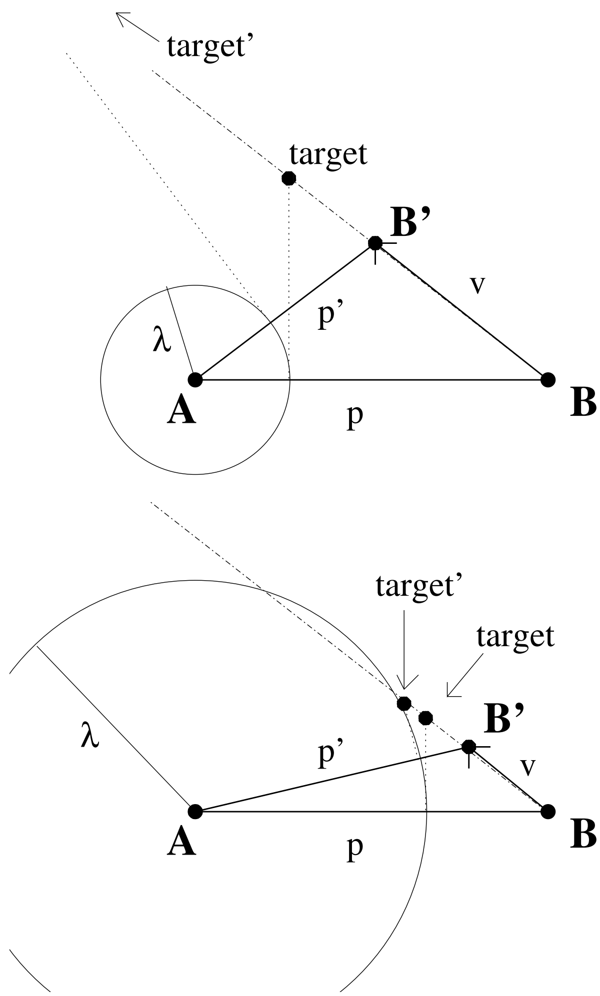


Figure 3. Motion of the target point over time; two cases.

Strengths and Weaknesses

The immediate-danger measure can be computed with only about 10 basic operations, even faster than the Close-Approach Model. Furthermore, it offers other useful information: The reciprocal of the measure is how much time remains until the airplanes play out whatever danger they face from each other.

This measure behaves correctly in two of the situations where the previous two measures required ugly fixes:

- If the two aircraft are on or very near a collision course, it acts as a countdown.
- If the airplanes are near in time to their closest approach, it gets large only if they are actually close to each other but remains small otherwise.

A flaw is that this measure always gives an immediate danger near zero when the airplanes have nearly identical velocities. As before, we'd like the danger in this case to be roughly inversely proportional to their separation. We can solve this problem by adding a term $c/|p|$ to the measure; unfortunately, doing so eliminates the nice relationship between this measure and the time left for the controller to act.

A potentially more serious problem is the inability of this algorithm to project far into the future. It detects almost no difference between, for example, two pairs of airplanes with the same relative velocities, the first of which are on course to collide in 5 min, and the second to pass with 1 mi of separation.

Testing the Models

In the situations of **Table 1**, all airplanes move at 480 knots (811 ft/s). Airplane A's initial position is at the origin.

Table 1.
Test situations.

Situation	A heading	B heading	B location (in ft)
1. Impending head-on collision	0°	180°	(6000, 0)
2. Impending oblique collision	60°	120°	(3000, 0)
3. Tailgating	0°	0°	(2400, 0)
4. Flying alongside	0°	0°	(0, 2400)
5. Same point, nearby time	0°	90°	(2400, −3200)
6. Same point, nearby time	0°	120°	(4400, −2100)
7. Passing at a distance	0°	180°	(18000, −6000)
8. Far-future head-on collision	0°	180°	(600000, 0)
9. Flying parallel	0°	0°	(0, 18000)
10. Right angles	0°	270°	(18000, 0)
11. Receding	0°	180°	(0, 6000)
12. Receding	120°	60°	(3000, 0)
13. Receding	180°	0°	(6000, 0)

We use our intuition and all of our models (except the Space-Time Model, for which we could not find appropriate constants in the allotted time) to rank the immediate dangers presented by each situation. With the Probabilistic Model, we use the metric

$$\text{danger} = \frac{\text{collisions}}{10000 \text{ trials}} + \frac{1}{20} \cdot \frac{\text{near misses}}{10000 \text{ trials}}.$$

The Close-Approach Model uses $c_1 = 50$ ft, $c_2 = 5$ s, and $c_3 = .05$ Hz^2.

Table 2.

Rankings of dangers of situations.

Situation	Intuition	Triv	Prob	Close-App	Logarithmic
1	1	2	1	2	3.5
2	2	2	2	1	3.5
3	3	9.5	4	3.5	10.5
4	4	9.5	7	3.5	10.5
5	5.5	4.5	5	5	2
6	5.5	4.5	3	6	1
7	8.5	9.5	11	9	5
8	7	2	6	10	7
9	8.5	9.5	8	8	10.5
10	10	9.5	11	7	6
11	12	9.5	11	12	10.5
12	12	9.5	11	12	10.5
13	12	9.5	11	12	10.5

Recommendations

Which Danger Model Should We Use?

The Probabilistic and the Close-Approach Models match well our intuitive rankings, though not particularly each other. The Trivial Model and the Logarithmic Derivative Model both compare much less favorably.

The Close-Approach Model agrees with our intuition almost exactly, except for switching the rankings of situations 8 and 10. As we expect very little eventual danger from situation 10, and very little immediate danger from situation 8, this is perhaps not so bad. The Space-Time Model would probably do even better: It would certainly rank situation 3 as more dangerous than situation 4 (agreeing with our intuition), and, as it is closely related to the Close-Approach Model, might well rank all the other situations identically.

We recommend the Close-Approach Model.

How Close Is Too Close?

In the Close-Approach Model, there is a natural map from the eventual danger to distance (by taking the reciprocal). Mahalingam [1999, 26–27] argues that airplanes should be horizontally separated by 3 nautical miles. But is she right?

Assume that each airplane has velocity v ft/s and that each can turn z radians/s. We find the distance d (in ft) at which each airplane must start turning to ensure that the aircraft do not pass within x feet of each other at any time.

Each turn (assuming constant turning rate) forms an arc of a circle of radius r. Since the length of an arc subtended by an angle θ is given by $s = r\theta$, we take the derivative to obtain

$$r = \frac{ds/dt}{d\theta/dt} = v/z.$$

Next, we note that x (the shortest distance between the two circles) is equal to the sum of the distance k between the centers of the two circles and the two radii: $k = x + 2v/z$ (**Figure 4**).

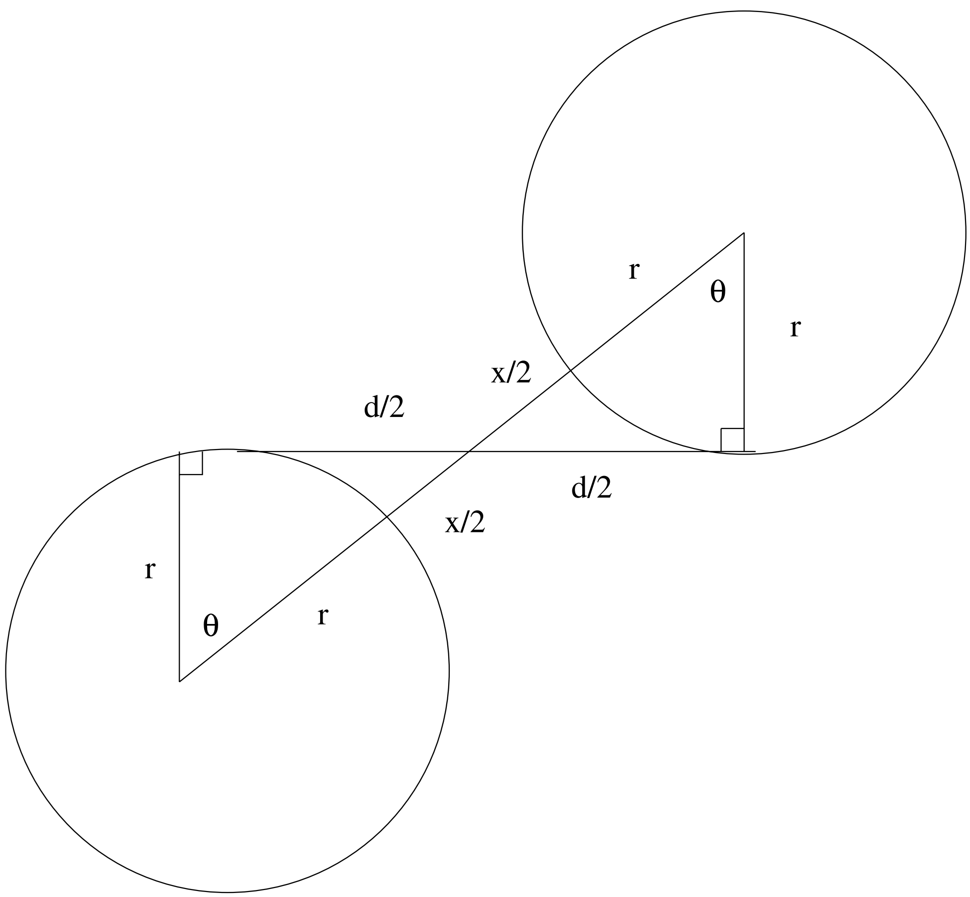


Figure 4. Analysis of avoidance of head-on collision.

Finally, we observe that the initial line of flight of the two airplanes is tangent to both circles and hence two right triangles are formed. In each triangle, the lengths of the sides are $r = v/z$, $d/2$, and hypotenuse $k/2 = v/z + x/2$. We apply the Pythagorean theorem to obtain $(d/2)^2 = (v/z + x/2)^2 - (v/z)^2$, so that $d = \sqrt{x^2 + 4xv/z}$.

We consider two airplanes with velocity 480 knots (811 ft/s) and turning rate $3^\circ = \pi/60$ radians/s (a standard "two-minute turn"). To avoid a near miss, the centers of the airplanes must be 700 ft apart. We calculate d = 6,621 ft = 1.09 nautical miles. So, both pilots must start a turn at a distance of at least 1.09 nautical miles apart, and the controller must identify the problem and communicate to the pilots before this point. Assuming a maximum delay of 15 s between the controller's discovery of the problem and the pilots' response gives a "safety distance" of about 5 nautical miles, or 19 s.

Measuring Complexity

To measure the complexity of the workload faced by an air traffic controller (ATC), we need a basic understanding of the tasks performed by the ATC and how ability to perform these tasks is affected by the number of airplanes in the airspace sector. We present the following algorithm as a model for the decision process of an ATC in detecting and solving conflicts. The order of the steps is drawn from a synthesis by Endsley and Rodgers [1994] of reports on the factors identified by experienced air traffic controllers as relevant to conflict prevention.

ATC Decision Algorithm

1. Scan the radar screen (and other sources of information) for airplanes located close to each other or currently at a safe distance but whose projected paths cross.

2. If a pair/group of airplanes at a given time instant appear close to each other, evaluate velocity and heading information to determine whether the airplanes will move to within a minimum separation distance of each other within the "near future" (2 min?).

3. If a potential conflict is detected, scan for other more pressing conflicts.

4. If there are no more-urgent conflicts, alert the pilots of the airplanes detected in Step 2 and formulate alternative routes for them.

5. Assess whether the alternative routes will cause conflicts with projected routes of nearby aircraft.

6. If the alternative routes will cause conflicts, reformulate other alternatives.

7. If there are no impending conflicts, or if the most recent conflicts have been resolved successfully, then take care of other tasks.

8. When the items in Step 7 have been adequately dealt with, return to Step 1.

Complexity of Step 1

For n airplanes in the airspace, there are $\binom{n}{2} = n(n-1)/2$ pairs, so this operation also has complexity of order $O(n^2)$.[1] More realistically, we could divide airplanes into clusters and analyze the complexity of each cluster, though clusters are not completely independent of one another.

Complexity of Steps 2–6

We incorporate the danger presented by each aircraft pair into a single danger metric D for the airspace by summing the dangers of the individual pairs. It is useful first to assess each danger against a threshold level for ATC intervention; then the measure of danger D for the airspace becomes the number of interventions that must be made.

For each pair of endangered airplanes, the ATC must resolve the situation while making sure that the solution does not conflict with the constraints of any previously solved conflicting pair. Thus, each conflict constrains the choices to resolve each of the other conflicts; in some cases, after the first k conflicts are solved, no solution for conflict $k+1$ may exist under the constraints, hence backtracking may be necessary. In other words, this problem is a form of the *general constraint satisfaction problem*, which is known to be NP-complete [Vardi 1999]. We guess that the worst-case complexity of our problem varies exponentially with the number of pairs: $O(k^{n(n-1)/2})$, for some constant k.

If the altitude of the airplanes cannot be changed, the ATC can either tell both pilots to bank to the right (from their point of view) or to the left. So, for every pair of aircraft in conflict, there are two possible solutions, hence 2^D possible solutions for a system with D conflicts.

We must also take into account the decisions of the ATC in Steps 4 through 6, estimating how many operations are involved. For an airspace divided into C clusters, with n_i airplanes in cluster i, there are $n_i - 2$ other airplanes to consider in resolving a conflict in cluster i. Thus there are (approximately) $2(n_i - 2)(D(i))$ interactions that the ATC must consider in a given cluster i. The number of interactions added by this measure does not change the complexity of Step 2, $O(k^{n(n-1)/2})$.

[1] EDITOR'S NOTE: Several collision-detecting algorithms are known to be more efficient in practice than $\mathcal{O}\left(n^2\right)$ without having theoretical guarantees. For references, see Eppstein and Erickson [1999]. They also give an algorithm to solve the problem in time $\mathcal{O}\left(n^{0.6897}\right)$ per collision using space $\mathcal{O}\left(n^{1.6897}\right)$, by means of ray shooting structures.

Additional Factors in Workload Complexity

Complexity is also affected by factors such as the rate of airplanes entering and exiting the airspace, the volume of the airspace, and the presence of additional software tools.

The complexity of airplanes entering and exiting is linear in the total number.

Many of the operational errors by ATCs result from ignoring secondary conflicts for too long [Endsley and Rogers 1997], so the potential for accidents is higher for more airplanes per unit volume.

Software could identify conflicts and order them by danger level, thereby reducing the complexity in Step 1; nevertheless, the primary complexity comes from solving conflicts once they arise ($O(n^2)$ for Step 1 vs. $O(2^{n(n-1)/2})$ for Step 2). So programs to detect conflicts do not combat the primary complexity faced by an ATC, and they could cause an ATC to take a more passive attitude in searching for potential conflicts not identified by the software. In other words, software could worsen the problem of ignoring secondary conflicts for too long. *Programs designed to aid an ATC in identifying conflicts should be designed as guides to the ATC's judgment rather than as automation of ATC functions*.

References

Airbus Industries. 2000. `http://www.airbus.com/imperial.html` .

Denker, John S. 2000. See How It Flies. `http://www.monmouth.com/~jsd/how/htm/maneuver.html` .

Endsley, M., and M. Rodgers. 1994. Situation awareness information requirements for en route air traffic control. DOT/FAA/AM-94/27.

________ . 1997. Distribution of attention, situation awareness, and workload in a passive air traffic control task. DOT/FAA/AM-97/13.

Eppstein, David, and Jeff Erickson. 1999. Raising roofs, crashing cycles, and playing pool: Applications of a data structure for finding pairwise interactions. Extended abstract at *Proceedings of the 14th Annual ACM Symposium on Computational Geometry* (1998), 58–67. *Discrete and Computational Geometry* 22 (4): 569–592. `\http://www.uiuc.edu/ph/www/jeffe/pubs/cycles.html` .

Federal Aviation Administration. 2000. Welcome to the Office of System Safety. `http://www.asy.faa.gov/asy_internet/safety_data/default.htm` .

Kremer, H., et al. 1997. Probabilistic vs. geometric conflict probing. National Aerospace Laboratory NLR. Amsterdam.

Lauderman, L.V., et al. 1998. Dynamic density: An air traffic management metric. NASA, NASA/TM-1998-112226.

Mahalingam, Sandra. 1999. *Air Traffic Control*. New Delhi: Kaveri Books.

MAICA: Modeling and analysis of the impact of changes in ATM. *Transport Research* #71. Belgium.

MICA: MET improvement for controller aids. *Transport Research* #72. Belgium.

Rodgers, M., et al. 1998. The relationship of sector characteristics to operational errors. DOT/FAA/AM-98/14.

Vardi, M. 1999. The descriptive complexity of constraint satisfaction. Implicit Computational Complexity 1999. `http://www.cs.indiana.edu/icc99/vardi.html`.

Wiener, E., and D. Nagel. 1988. *Human Factors in Aviation*. San Diego: Academic Press.

The Safe Distance Between Airplanes and the Complexity of an Airspace Sector

Finale Doshi
Rebecca Lessem
David Mooney
Governor's School
Richmond, VA

Advisor: Crista Hamilton

Summary

We determine the minimum safe spacing between aircraft and also the complexity of the air traffic control system.

Taking into account the vortex that a leading plane leaves in its wake, the distance between the tail of one plane and the nose of the next plane should be at least 5.5 km or 3.4 mi. The minimum spacing between adjacent planes either to the side, above, or below should be at least 730 m or 0.45 miles. These distances were calculated using Bernoulli's principle, which states that the internal pressure of a fluid (such as air) decreases when its speed increases. Because the speed of a plane is very high, the pressure around the wings is low. The change in pressure associated with the Bernoulli factor, applied over the facing surface area, results in a force pushing the planes together; the force may alter the plane's flight pattern.

Finally, if two planes are heading towards each other, there must be enough space between them to perform evasive maneuvers. We find that 12 s is required; at normal flight speed, this translates to 2.9 km or 1.8 mi.

We define complexity of an airspace sector as the probability of a conflict occurring during a given period of time. To determine complexity, we assume that sectors are rectangular solids and that planes fly either in parallel or antiparallel directions. We calculate the probability that a plane enters a sector either too soon after another plane, or that two planes enter the same airway going in antiparallel directions.

The UMAP Journal 21 (3) (2000) 257–267.

The weaknesses of this model include that all planes are assumed to be Boeing 767s. This model also does not take into account weather changes and multiple conflicts.

The strengths of this modeling include allowing for passenger safety while slightly shortening the Federal Aviation Administration (FAA) minimum distances, thereby increasing airspace capacity. The complexity model accounts for stress; stability analysis shows that a small change in the environment does not drastically change the model.

Background

According to current FAA separation guidelines, an aircraft must maintain a separation of 5 mi behind the plane and 2 mi adjacent to the plane [Gilbert 1973, 36–37].

Numerous benefits could come from reducing separation standards. Primarily, air space capacity would increase. Delays would be reduced because planes would not have to wait for an open airway. Finally, fuel costs would decrease because planes would be rerouted less frequently, with fewer delays.

Questions remain, however, as to whether other bottlenecks would mitigate these benefits. Potential conflicts can occur in one of two areas:

- Over 75% of collisions occur within the terminal area—the space within 30 mi of an airport—because traffic is dense and constantly changing.
- Other conflicts occur en-route; most en-route airplanes have a constant altitude and velocity [Gilbert 1973, 91].

Our model concentrates on en-route traffic.

The purpose of the air traffic control system is to avoid collisions. Avoidance depends on two basic sources of air traffic information:

- Air-derived collision avoidance involves either the visual detection of a conflict by the pilot or the radar detection of air disturbances around the plane.
- Ground-derived collision avoidance uses radar from ground-based radar antennae. Controllers monitor radar to detect potential conflicts and contact the pilots involved to give them new courses.

Ground-derived avoidance is the primary tool to maintain minimum spacing between planes. The controller of a certain air space gives continuous and detailed instructions to the pilot as to flight parameters that should be taken in the airspace, including heading and altitude. When all aircraft operate under the Ground Collision Avoidance System (GCAS), there is an extremely low rate of mid-air collisions [Collins 1977, 123].

The type of radar most commonly used by GCAS for en-route traffic monitoring is Air-Route Surveillance Radar (ARSR), long-range radar with a range of 200 mi and an altitude range of 40,000 ft. It has a slower rotation (3–6 rpm) than short-range radar (10–30 rpm); because the rotation is slower, accuracy and resolution are not as high [Gilbert 1973, 40].

Other disadvantages in the current radar system are [Federal Aviation Administration 1997, 4.2]:

- ARSR lacks sufficient low-altitude coverage because traffic is concentrated at higher altitudes.
- Radar equipment can be unreliable or malfunction.
- Blind spots exist in the radar pattern, behind large objects and other planes.
- Radar cannot differentiate between targets within 3° of each other from the radar antenna; these objects blur together on the screen.

The capacity of the airspace is limited by the minimum spacing between planes and also by the size of the workload placed on the controller team.

Airspace is broken down into *sectors*; one controller team manages each sector. The controllers must maintain radio contact with each aircraft located in the sector, and they must identify each aircraft on the radar screen. Each aircraft must be assigned a "travel plan" with a vector, heading, and altitude. Controllers must maintain constant surveillance on each aircraft flight pattern to identify potential conflicts. If the number of aircraft in a sector increases, more work is required by the controllers and more separation between planes is necessary to ensure that controllers spot potential conflicts. Controllers must also "transfer" planes that are exiting to a neighboring sector [Federal Aviation Administration 1997, 4.2].

Dividing the airspace into a greater number of smaller sectors would ease the workload of controllers, but the increased work in "sector transfers" and the increased cost in added controller teams and radar would reduce efficiency.

Assumptions in the Model

- All planes behave like Boeing 767s.
- All planes are considered cylindrical rings for physics calculations.
- Weather is not a factor in the safe distance between planes.
- All planes are en route; they are not landing or taking off.
- All planes are flying at about 35,000 ft.
- All planes have a constant speed of 857 km/hour.
- There is no energy loss due to friction.

Table 1.
Symbols in the model (with standard SI-MKS units).

Symbol	Meaning
a	Acceleration
α	Angular acceleration
$C(t)$	Number of planes in crucial beginning area of sector at time t
d	Distance between planes
$\Delta\theta$	Angular rotation
h	Height of sector
I	Rotational inertia
L	Length of sector
ω	Angular velocity
P	Pressure of air
$P(t)$	Probability of conflicts at time t
$R(t)$	Number of planes entering sector at time t
ρ	Density of air
t	Time
v	Velocity
w	Width of sector

Model Development

The model examines the distances that are required between planes from the front, back, above, below, and laterally; different forces and factors account for the different distances required.

When two planes in opposite directions approach each other, we assume that one of the planes descends to avoid a collision. Knowing that it takes 12 s to avoid a collision, we can find the minimum distances. Each plane creates a pair of vortices, areas of strong turbulence, that extend outward and around from the wingtips. A vortex affects the safe distance behind the plane (point c in **Figure 1**); the vortex from a large plane is large enough to damage seriously another plane that follows too closely.

The other important factor is Bernoulli's Principle, which states that the internal pressure of a fluid (liquid or gas) decreases at points where the speed of the fluid increases. The moving air around the wing reduces the air pressure and would cause a nearby plane to accelerate towards the first plane. The force affecting planes above, below, and to the sides (points d, b, e, and f) comes from a combination of the Bernoulli and vortex forces. Enough distance must be allowed between planes to overcome it.

Distance Required in Front of Plane

According to 241 Air Traffic Control Squadron (ATCS) [1999], 12 s is needed to steer clear of another object: 6 s for the controller to radio to the pilot, 4 s for the pilot to start the maneuver, and 2 s to gain enough space to clear. At a speed of 238 m/s, the corresponding distance is 2.9 km.

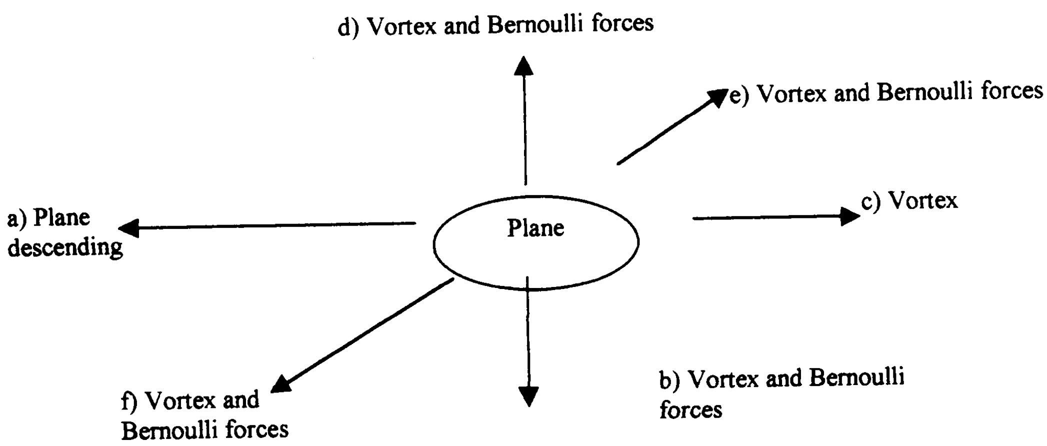


Figure 1. Plane and associated forces.

Distance Required Behind Plane

The minimum safe distance below and behind the plane is determined by the size of the vortex behind the plane. Each wing develops a vortex of air approximately 15.0 m in diameter, spinning at 42.7 m/s. The vortex sinks at 2.03 m/s until it is approximately 244 m below the level of the plane; it is usually 9,250 m long. This vortex is a great danger to following planes because it can cause them to roll [241 ATCS, 1999].

Because the Boeing 767 is 48.5 m long, and the vortex has a diameter of 15.0 m, a column of air with volume 8,850 m^3 acts upon a plane flying in a vortex from another plane. The air density at 10.7 km above the ground (cruising altitude for Boeing 767s) is 0.380 kg/m^3. Hence, the mass of air acting on the plane is 3,360 kg if the plane is flying into the vortex.

We assume that all of the angular momentum of the air is transferred to the plane:

$$I_{\text{air}}\,\omega_{\text{air}} = I_{\text{plane}}\,\omega_{\text{plane}}, \tag{1}$$

where I is rotational inertia and ω is angular velocity.

The air is a spinning disk, whose rotational inertia is given by

$$I_{\text{air}} = \frac{mr^2}{2} = 0.5 \times 3{,}360 \times (7.502)^2 = 9.77 \times 10^4 \text{ kg} \cdot \text{m}^2.$$

The angular velocity of the air at given distance d from the plane can be found by using the properties of the vortex. As it leaves the plane, winds near the edge of the vortex measure at 45.72 m/s. The circumference of the vortex is 47.1 m; it therefore takes 1.05 s for the air to travel through one rotation, which means that the initial angular velocity is 6.00 rad/s. The vortex disappears in 9,250 m; because the plane flies at 238 m/s, this is 38.8 s after the vortex was created. The angular deceleration, then, is found using the formula

$$\alpha = \frac{\omega_f - \omega_i}{t} = \frac{0 - 6.00}{38.8} = -0.155 \text{ m/s}^2.$$

The air decelerates at this rate. The plane moves at 238 m/s, so it will take $d/238$ s to travel distance d. The equation for angular velocity at distance d becomes:

$$\alpha = \frac{\omega_f - \omega_i}{t} = \frac{\omega_f - 6.00}{d/238} = -0.155\ \mathrm{m/s^2}.$$

We solve to find angular velocity ω_f in terms of d:

$$\omega_f = (-6.49 \times 10^{-4})d + 6.00\ \mathrm{rad/s}.$$

Next, we consider the angular velocity and rotational inertia. We assume that the plane spins around its central axis. Because most of the plane's mass is located to the outside, we assume that the plane is a rotating ring. Using the mass and radius of the Boeing 767, we find the rotational inertia:

$$I = mr^2 = (156{,}500)(2.85)^2 = 1.27 \times 10^6\ \mathrm{kg \cdot m^2}.$$

This calculation assumes that the plane is flying into the vortex, not across it. Across the vortex, its rotational inertia would be that of a pivoting rod, given by:

$$I = \frac{ml^2}{3} = \frac{(156{,}500)(48.5)^2}{3} = 1.23 \times 10^8\ \mathrm{kg \cdot m^2}.$$

This is far greater inertia than for the plane flying into the vortex. Less inertia means that it takes less force to turn the plane, meaning that the situation is more dangerous. Therefore, to determine the safe distance, we consider further only the approach from behind.

The plane starts with zero angular velocity. We assume that it cannot turn more than 5° (0.0873 radians) in the crossing period (0.938 s) without discomfort and loss of control. Further, we assume that the plane is climbing at an angle of 5° or steeper as it goes through the vortex. At this angle, the plane, traveling at 238 m/s, would be in the vortex for no more than 0.938 s. The final angular velocity of the plane is found using these data and the equations

$$\Delta\theta = \omega_i t + \alpha t^2/2, \quad 0.0873 = 0 + \alpha \times (0.938)^2/2, \quad \alpha = 0.198\ \mathrm{rad/s^2};$$
$$\omega_f^2 = \omega_i^2 + 2\alpha\Delta\theta = 0 + 2 \times 0.198 \times 0.0873, \quad \omega_f = 0.186\ \mathrm{rad/s}.$$

Substituting these values into **(1)** gives

$$(9.77 \times 10^4)[(-6.49 \times 10^{-4})d + 6.00] = (1.27 \times 10^6)(0.186),$$
$$d = 5{,}510\ \mathrm{m}.$$

Adding 152 m for radar uncertainty, we get an unsafe zone 5.7 km long behind a plane.

Distance Required Vertically and Laterally

We use Bernoulli's Principle to find the minimum distance on the sides and up and down between planes. The distance between the two planes is d, the initial velocity of the vortex is 45.7 m/s, and the acceleration of the air is the same as the acceleration of the vortex, -1.766 m/s^2.

We apply the equation

$$v_f^2 = v_i^2 + 2ad = (45.7)^2 + 2(-1.77)d = 2.09 \times 10^3 - 3.53d.$$

Next, to determine the change in pressure, we apply Bernoulli's equation governing fluids:

$$P_i + \frac{\rho_i v_i^2}{2} = P_f + \frac{\rho_f v_f^2}{2},$$

where ρ is air density. The initial velocity is zero, and at 35,000 ft the density is 0.380 kg/m^3 and the atmospheric pressure is 2,340 Pa:

$$2.34 \times 10^3 = P_f + \frac{(.380)(2.09 \times 10^3 - 3.53d)}{2}, \qquad P_f = 1.95 \times 10^3 + 0.670d \text{ Pa}.$$

The change in pressure is

$$2.34 \times 10^3 - (1.95 \times 10^3 + 0.670d) = 397 - 0.670d.$$

Since pressure equals force divided by area, the force exerted can be found by using the surface area of the plane, which varies depending on whether the sides or the top and bottom are being considered.

Sides

The length of a Boeing 767 is 28.5 m and the width is 5.7 m. Assuming that the plane is a cylinder and half of it is facing the side, the surface area affected is $(28.5)(5.7)\pi/2 = 434$ m^2. (This is an overestimate, since we neglect the curvature of the body of the plane.) Since pressure = force/area, we have $397 - 0.670d =$ force/434 and the force is $1.72 \times 10^5 - 291d$. We assume that the maximum lateral acceleration, before either the ride becomes too turbulent or the plane loses control, is 0.1 m/s^2. Newton's second law (force = mass $\times$ acceleration) becomes $1.72 \times 10^5 - 291d = (1.57 \times 10^5)(0.1)$, leading to $d =$ 538 m.

The total distance on the sides now needs to be calculated:

$$\begin{aligned}\text{Distance} &= 538 + \text{Vortex Width} + .5(\text{wingspan}) + \text{Uncertainty factor for radar} \\ &= 538 + 15 + 47.6/2 + 152 = 729 \text{ m}.\end{aligned}$$

Therefore, 729 m must be allowed on each side of the plane.

Above and Below

For the vertical viewpoint, we have

$$\text{Surface area} = \text{half cylinder} + \text{wing area} = (48.5)(5.7)\pi/2 + 283 = 717\text{m}^2.$$

The pressure, calculated previously, is $396.8 - 0.670d$. We have

$$\begin{aligned} \text{Pressure} &= \text{Force}/\text{Area}, \\ 397 - .670d &= \text{Force}/717, \\ \text{Force} &= 2.85 \times 10^5 - 481d. \end{aligned}$$

Using Newton's Second Law as before, we have

$$F = ma, \qquad 2.85 \times 10^5 - 481d = (1.57 \times 10^5)(0.1), \qquad d = 559 \text{ m}.$$

The total vertical distance is

$$\text{Distance} = 559 + \text{Vortex Width} + \text{Uncertainty factor} = 559 + 15 + 152 = 727 \text{ m}.$$

Therefore, for safety, a plane needs 729 m on each side and 727 m above and below.

Complexity

Complexity, from a workload perspective, we define to be the probability of a conflict—and therefore the need for an air traffic controller (ATC) to intervene—in a certain period of time. A high likelihood of conflict in a short period of time means a lot of potential stress for the ATC. In addition, to accommodate time to recover from stress, the 5 min before the time being considered is also included in the definition of complexity.

We first find the probability of a conflict at a certain point in time. We assume that sectors are rectangular solids ($L \times w \times h$, in m). We divide a sector into *blocks* 727 m tall and 729 m wide, the minimum for Boeing 767 planes to be apart.

We assume that planes fly in either parallel or antiparallel directions. There are two possibilities for conflict.

- A plane enters a block too soon after another plane. The safe distance between planes following each other—the sum of the minimum distance in front of and behind a plane, 8,380 m—should not be violated; so planes entering a block should be 35.2 s apart.
- Two planes enter the same block from opposite ends, in antiparallel directions.

We assume that planes enter a block according to a random process, so that entries to the block are independent.[1]

[1] EDITOR'S NOTE: The presentation here has been adapted slightly from the authors'. The problem that they consider involves the composition of two independent *renewal processes*, one at each end of the block. Consider a block *available* if there is no plane in it; the limiting availability A of a block, as $t \to \infty$, depends only on the mean time between arrivals and the mean time to pass through the block [Trivedi 1982, 297–301 and 305, Problem 1]. The same kind of rectangular block model was used to calculate the probability of a mid-air collision over the North Atlantic [Machol 1975]; the editor thanks Antony Unwin of Universität Augsburg, Germany, for supplying this reference. See also the update at Machol [1995].

Let $m(t)$ be the probability density function for entry of a plane into a block, divided equally between the two directions.

Suppose that a plane enters the block at time t.

- The probability that another plane enters the block in the same direction as the first plane but too close behind it is

$$P_1(t) = \frac{1}{2} \int_t^{t+35.2} m(t)\, dt.$$

- The probability that another plane enters the block from the opposite direction while the first plane is in the block is

$$P_2(t) = \frac{1}{2} \int_t^{t+L/238} m(t)\, dt.$$

Hence, the conditional probability of conflict in a block, given entry of a plane into the block at time t, is the sum of the two probabilities:

$$P(\text{conflict} \mid \text{plane enters at } t) = \frac{1}{2} \left[\int_t^{t+35.2} m(t)\, dt + \int_t^{t+L/238} m(t)\, dt. \right].$$

The unconditional probability density of conflict, assuming independent arrivals of planes in the block, is

$$\begin{aligned} P(t) &= P(\text{conflict} \mid \text{plane enters at } t)\, P(\text{plane enters at } t) \\ &= P(\text{conflict} \mid \text{plane enters at } t)\, m(t). \end{aligned}$$

The most important component of complexity is the probability of a conflict; the most complex situation is a high likelihood of conflict over a sustained period of time.

However, a situation is more complex if the ATC is already stressed from previous problems, so in our definition of complexity for a block over an interval we add, weighted at 10%, the probability of a conflict during the 5 min (300 s) preceding:

$$\text{Complexity for a block} = P(t) + 0.1 \int_{t_1-300}^{t_1} P(t)\, dt.$$

While software to alert controllers of potential conflicts would add to the safety of flight, it would also add to the conflict. Most models and distance estimates, including this model, tend to overestimate the safe distance between planes. Even if this were remedied, there is also the uncertainty of the exact position of the plane due to the imprecision of the radar, which would cause the software to alert the ATC even when no conflict was likely to occur. This would add to the complexity, because the controller would have more potential conflicts through which to sift. However, this added complexity would add to the safety because most possible conflicts would receive ample warning.

Stability

We did a stability analysis in regard to changes in the mass of the planes, the velocity of the vortex, and the pilot's reaction time. The mass of the planes or the velocity of the vortex would be different for a different type of plane, and pilots' reaction times may vary.

With the lateral Bernoulli forces, if the initial velocity of the vortex is changed by 2.2%, the distance is changed by 4.8%. If the mass is changed by 6.4%, the distance is changed by 0.5%. With the vertical Bernoulli forces, if the initial velocity of the vortex is changed by 2.2%, the distance is changed by 4.5%.

The effect of the changes in the pilot's reaction time was analyzed in regard to the distance in front of the plane. If the time is changed by 8.3%, the distance also changes by 8.3%. Therefore, this model is stable with regard to all of the variables tested.

Strengths and Weaknesses

The model provides for safety. Parameters, such as the mass of the plane, can be changed as appropriate. However, the model does not accommodate two planes of different models. The complexity section accounts for controller stress. The model's minimum spacings for aircraft, including taking into account 500 ft of radar uncertainty, reflect FAA guidelines reasonably well.

References

241 Air Traffic Control Squadron (ATCS). 1999. 241 ATCS Home Page. Modified December 1999. Accessed February 2000. `http://aw139.ang.af.mil/241atcs` .

Association for the Advancement of Amateur Aeronautics and Astronautics. 1999. Altitude Range: General. Modified November 1999. Accessed February 2000. `http://www.tinkertech.com/a5/outline/altrange.gen.html` .

Boeing Corporation. 2000. Boeing Home Page. Modified February 2000. Accessed February 2000. `http://www.boeing.com` .

Collins, Richard L. 1977. *Flying Safely*. New York: Delacorte Press.

Federal Aviation Administration. 1997. *Pilot's Handbook of Aeronautical Knowledge*. Washington, DC: Government Printing Office.

Giancoli, Douglas C. 1991. *Physics: Principles with Applications*. 3rd ed. Englewood Cliffs, NJ: Prentice Hall.

Gilbert, Glen A. 1973. *Air Traffic Control: The Uncrowded Sky*. Washington DC: Smithsonian Institution Press.

Machol, Robert E. 1975. Navigation standards over the North Atlantic. *Interfaces* 5 (No. 2, Pt. 2) (February 1975): 62–71.

________. 1995. Thirty years of modeling midair collisions. *Interfaces* 25 (5) (September-October 1995): 151–172.

Marks, Robert W. 1969. *The New Dictionary and Handbook of Aerospace*. New York: Frederick A. Praeger Publishing.

Shevell, Richard S. 1983. *Fundamentals of Flight*. Englewood Cliffs, NJ: Prentice Hall.

Trivedi, Kishor S. 1982. *Probability and Statistics with Reliability, Queueing, and Computer Science Applications*. Englewood Cliffs, NJ: Prentice-Hall.

________. 1998. *A Concept Paper for Separation Safety Modeling: An FAA/Eurocontrol Cooperative Effort on Air Traffic Modeling for Separation Standards*. Washington, DC: GPO.

Wesson, Robert, et al. 1981. *Scenarios for Evolution of Air Traffic Control*. Santa Monica, CA: Rand Corporation.

The Iron Laws of Air Traffic Control

Kevin Arnett
Jonathan S. Gibbs
John J. Horton
U.S. Military Academy
West Point, NY

Advisor: David Bailey

Introduction

We focus our analysis on two key system design specifications:

- how the Air Traffic Control System accomplishes its requirements (the safe routing of aircraft through a sector of airspace), and
- what computational and time demands are generated by the traffic load.

We develop a solution that takes into account factors such as knowledge of proposed flight paths, orientation of aircraft, and acceptable probability of an accident. We develop two separate models to analyze situations where in-flight conflicts arise.

- The first examines the position of aircraft through three-dimensional normal probability distributions to develop the likelihood of a collision.
- The second uses vector calculus and dynamics to develop real-time data on the likely trajectory of an aircraft, making no assumptions that the aircraft are flying along a predetermined path.

Two other models use analogies to other fields to provide metrics of complexity of the workload of the air traffic controller (ATC), one focusing on the inherent complexity of an airspace (analogous to fluid flow) and the other on number of aircraft.

Of our four models, we are able to validate only two, the probability distribution model for likelihood of collision and the airspace complexity model. The implementation and validation are straightforward and we omit them due to time considerations.

The UMAP Journal 21 (3) (2000) 269–282.

Assumptions

- All aircraft flight paths are filed with the FAA prior to departure and are available to all ATCs.
- We treat an aircraft as a sphere with a set radius, whose orientation is insignificant.
- Two distinct types of errors cause flight path deviation:
  - systematic error (pilot error, emergencies, failed equipment, etc.) and
  - random error (differences in weather, GPS signals, aircraft characteristics).
- Relativistic effects are negligible.
- The curvature of the Earth need not be accounted for explicitly (the ease with which transformation matrices can be designed precludes this from being of great import).

Requirement A

All aircraft are required by law to file flight plans with the Federal Aviation Administration. ATCs generally assume that an aircraft will follow its filed flight plan but are attentive to deviations. There is a duality in the ATC job: on one hand, ATCs plan as if everything will perform in a predictable manner, while they simultaneously must be vigilant in case things do not. The two different roles of the ATC are reflected in our model. Based on how an ATC monitors a sector, we break our model into two separate submodels:

- The Random Effects Model predicts potential conflicts between two aircraft based on flight path data and characteristics of the aircraft, weather, etc.
- The Contingency Model addresses routine real-time monitoring of aircraft. It makes no assumptions about aircraft following a given flight path. The model serves as an alert system for an ATC, warning of conflicts caused by systemic errors as they arise.

Both models run on real-time data, but the Contingency Model has a higher priority in terms of computing resources. The Random Effects Model is re-evaluated as fluctuations in velocity affect the flight path. The Contingency Model is more relevant in situations where airplanes are not well spaced and where there is a high probability for unplanned deviations in flight path. In contrast, the Random Effects Model is more suited to the operational planning of flight paths and the monitoring of major air corridors.

While certainly different, these two models both answer the question of what constitutes "too close," by examining situations that make a collision likely and advising the ATC of the need for intervention.

It may seem that we should determine a safe separation between two planes, but this question is so related to orientation as to be meaningless. The better approach is to determine what flight paths and velocities will cause a collision.

Random Effects Model

We represent an expected path with a vector-valued function $r_p(t)$ such that

$$r_p(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k},$$

where t is time and x, y, z are functions that describe the aircraft's position. However, due to a variety of factors such as weather and instrumentation inaccuracies, the actual position is not fixed but rather is dependent on three random variables. We define this actual position vector as

$$r_p(t) = x'(t)\vec{i} + y'(t)\vec{j} + z'(t)\vec{k} = [x(t) + \epsilon_x]\vec{i} + [y(t) + \epsilon_y]\vec{j} + [z(t) + \epsilon_z]\vec{k},$$

where each of the three error terms has normal distribution centered at 0, $\epsilon_i \sim N(0, \sigma_i^2)$ for dimension i, where each dimension can have a different variance. We assume independence of the error terms. We identify values for the variances based on data from the FAA (see **Appendix A**). Essentially, this model describes a probability shell surrounding the aircraft (**Figure 1**).

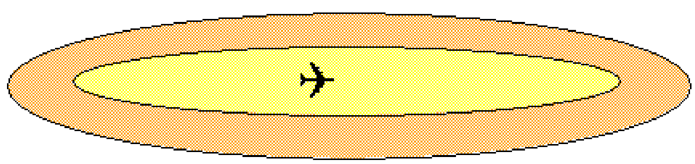


Figure 1. Probability shell for an aircraft in flight (not to scale).

The probability of being within Δ_i of a predicted location in dimension i is

$$P_i(\Delta_i) = \Phi(\Delta_i/\sigma_i) - \Phi(-\Delta_i/\sigma_i) = 2\Phi(\Delta_i/\sigma_i) - 1,$$

where Φ is the standard normal cumulative probability density function. Assuming independence of the error terms in the three directions, the probability that the plane will be within $(\Delta_x, \Delta_y, \Delta_z)$ of its projected position is

$$g(t, x', y', z') = P_x(\Delta_x)P_y(\Delta_y)P_z(\Delta_z).$$

Consider two aircraft with probability functions g and h, assumed independent. The probability that there will be a collision at a given point is the probability that both planes occupy that point at the same time, namely,

$$c(t, x', y', z') = g(t, x', y', z') \cdot h(t, x', y', z').$$

Thus, we determine the viability of two flight paths by examining the probabilities associated with each point at every instant. This is a very computationally intensive prospect, but there are methods that provide solutions in reasonable amounts of time.

Implementation and Validation

We wrote a program in Visual Basic for Applications on top of a Microsoft Excel spreadsheet. Flight data are placed in the worksheet, which the program uses as input. We used the program to validate the model for two scenarios:

- The two planes collide.
- The two planes cross paths but at different times.

The results of these simulations are presented and discussed in **Appendix B**.

Contingency Aircraft Tracking System

For aircraft deviating grossly from their flight plans, the Random Effects Model is not useful. The Contingency Aircraft Tracking System is designed to alert the ATC to any aircraft that could be on a collision course.

Using data collected on the positions of the aircraft, by either GPS or some other monitoring system, the system uses the path that the aircraft has been on to predict where it will be in the future. Those future positions are used to designate a sector of air space as off limits to other aircraft. When two or more aircraft are predicted to pass through the same sector, the ATC is alerted. This tracking system is based on several assumptions:

- Aircraft do not accelerate in the direction of travel while in the airspace.
- Aircraft turn at a constant normal acceleration or move in a straight path.
- The position locating systems are accurate and provide continuous updating.

Most major airports have the tracking ability described in the third assumption; continuous updating compensates for the other assumptions of zero tangential acceleration and constant normal acceleration.

To predict possible future positions of the aircraft, we use three past positions. We calculate the vectors from the first to the second and from the second to the third and determine the angle between them. If the angle is below a certain tolerance, we approximate the path of the aircraft by a straight line; otherwise, we approximate it by the arc of the circle defined by the three points. In the latter case, since the aircraft could also stop turning or turn less sharply, we generate a line for the path of the aircraft if it keeps its current velocity vector, tangent to the circle. Anywhere between the arc of the circle and the line—a planar, fin-shaped area—is a possible future position. This system approximates the fin with a triangle defined by the current aircraft position, the position the aircraft would occupy if it continued along the arc of the circle for the time step, and the position if it flies straight for the rest of the time step. [EDITOR'S NOTE: We omit the vector calculus details of the calculations involved.]

Requirement B

Applying vector calculus and multivariable analysis, we relate the characteristics of the flow of traffic in a sector to the complexity of the ATC's job in controlling the sector. By examining the influence of aircraft entering and exiting the sector in three different respects (instantaneous, over a time interval, and over a particular time of day), we refine the model.

Determining the Complexity Inherent in a Sector

The sector is defined by its size (boundaries) and by the objects that impact the flow of traffic through it. We assume that the sector extends from the ground upwards through all space (no ceiling) and is bounded by cylinder walls following the shape of the base of the sector (ground projection).

We model the airport as a vector field with the following properties:

- Aircraft are equally drawn from all points toward the location of the airport.
- The size of the airport determines the magnitude of the attraction impact on aircraft traffic.

The simple field to address these requirements is

$$A(x,y) = k_i \cdot \frac{(x_i - x)\vec{i} + (y_i - y)\vec{j}}{|(x_i - x)\vec{i} + (y_i - y)\vec{j}|}.$$

The field A_i for airport i points from any point in the plane (x, y) toward the airport location (x_i, y_i). Furthermore, we assign each vector a magnitude k_i, representing the impact of the airport on the traffic in the sector.

The properties for obstacle fields in the sector are:

- An obstacle's influence on a point is limited by the distance from the point to the center of the obstacle (obstacles create local effects).
- Physically larger obstacles impact aircraft farther from their centers than smaller obstacles do (Dallas/Ft. Worth has a greater influence on traffic than a municipal landing strip).
- The impact of the obstacle on traffic is related to properties like permanence, physical height, and the way that it impacts air traffic (a small town should have a lesser impact than a similarly sized downtown of skyscrapers).

We use the following function for obstacle intensity:

$$O_i(x, y) = -h_i \exp\left[\frac{(x - q_i)^2 + (y - w_i)^2}{-l_i}\right].$$

Points more distant from the center of the obstacle (q_i, w_i) get values closer to zero, and the variable l_i reflects how large (geographically) the obstacle is. Our final property is upheld by the use of h_i, which determines the relative impact of the obstacle. The negative sign ensures that the object repulses traffic.

The impact of a number of obstacles is just the sum O_{TOT} of their impacts, and we get the vector field for the obstacles by taking the gradient $B(x, y) = \nabla O_{\text{TOT}}$ of the total impact. This creates a field in which traffic tends to flow away from obstacles in the radial direction.

We then combine the obstacle and airport fields into a single flow field through simple addition.

Total Complexity

We characterize the complexity of the flow field that we have created. The flow of aircraft through a sector is analogous to the flow of fluid particles during bulk flow. In fluid mechanics, a laminar flow is marked by smoothness and predictability and is irrotational (no eddies). Turbulent flow, far more difficult to analyze, is choppier, less predictable than laminar flow, and rotational. Turbulent flow through the sector serves as our model for high complexity, while laminar flow is analogous to a sector with very low complexity.

The curl of a vector field measures the level of rotationality of a flow. We evaluate the magnitude of the curl of the vector field at every point and use this as the measurement of the total complexity of a sector.

Complexity by Number of Aircraft

We examine the impact of traffic volume on complexity of the workload for ATCs. We delineate three separate components:

- instantaneous complexity,
- complexity over a time interval, and
- complexity over a particular time of day.

From most demanding to least demanding, the tasks of an ATC are:

- adjust a plane's trajectory to avoid a potential conflict or collision,
- create a minimum spanning tree that highlights the critical relationships among aircraft,
- calculate the distance between aircraft,

- receive and record data from each aircraft, and
- communicate with ATCs in adjacent sectors (hand off aircraft to one another).

Definition of Complexity

We propose that the complexity of monitoring a sector can be defined similarly to the time complexity of an algorithm, in terms of the number of reference functions required to produce the correct output.

Instantaneous Complexity

We assume that something similar to our flight plan validation model is used by the ATC, screening potential conflicts long before the concerned aircraft enter the sector. The real-time complexity of the ATC's workload is then related solely to the need for corrections and the management of aircraft in the sector. The instantaneous case handles deciding if corrections are necessary, which requires examining the relationship between every pair of aircraft. But checking all of the $\binom{n}{2} \sim \mathcal{O}\left(n^2\right)$ distances between pairs is inefficient; a human operator visually inspecting graphical output should be able catch dangerous interactions between aircraft, at least in simple or routine situations. To do so requires the ATC to look at only $n-1$ interactions, examining only the distance between an aircraft and its nearest neighbor. In more complex situations, we would need a better method for determining interactions of concern.

An improved process for extremely vexing scenarios would be to employ a minimum spanning tree algorithm to determine the "edges" of interest, namely, the distance between an aircraft and its closest neighbors. By the definition of a minimum spanning tree, all airplanes that are very close together are connected by an edge, whereas those that are relatively far away from each other are not. Minimum spanning tree algorithms, such as Prim's or Kruskal's algorithms, have a complexity of $\mathcal{O}\left(n^2\right)$. At first glance, this would not appear to have an advantage over checking the distances between all pairs. But the minimum spanning tree would not have to be determined at every iteration; the tree could be reused for some number of time steps without significant loss of accuracy. Hence, we conjecture that the instantaneous time complexity of monitoring n aircraft falls between $\mathcal{O}(n)$ and $\mathcal{O}\left(n^2\right)$.[1]

Time Interval Complexity

We examine how complexity is related to the number of aircraft passing through the sector over a given interval of time. The difference between this

[1] EDITOR'S NOTE: Several collision-detecting algorithms are known to be more efficient in practice than $\mathcal{O}\left(n^2\right)$ without having theoretical guarantees. For references, see Eppstein and Erickson [1999]. They also give an algorithm to solve the problem in time $\mathcal{O}\left(n^{0.6897}\right)$ per collision using space $\mathcal{O}\left(n^{1.6897}\right)$, by means of ray shooting structures.

scenario and the instantaneous example is that there are now additional tasks beyond simply monitoring positions:

- receive and record updated aircraft positions,
- deal with aircraft entering the sector,
- hand off aircraft leaving the sector, and
- issue flight adjustments.

As before, we are not interested in how long these tasks take, or even how many are performed; our concern is how the number of these operations that must be performed is related to the number n of planes passing through the sector.

- **Updates:** To update the positions of the airplanes in the sector, n transmissions from the aircraft must be received and recorded.
- **Flight Adjustments:** We assume that any given aircraft has a probability p of requiring a path readjustment at a given instant. It is reasonable to conjecture that this probability increases by (at least) some constant amount for each aircraft added to the system, so we postulate that $p = cn$, where c is some constant. The probability that a given plane will require an adjustment, multiplied by the total number of planes in the system, provides an estimated readjustment workload for the ATC of $W = cn^2$.
- **Handoffs and Receptions:** For a plane entering the sector, its relationship with other aircraft already in the sector must be evaluated. Thus, entering aircraft require distance calculations. Aircraft leaving the sector require that the ATC send a radio transmission to the adjacent sector's ATC. An exiting aircraft is no longer tracked, so complexity decreases. We ignore the cost of handing an aircraft off, as it is in terms of cost.
- **Total Analysis:** The complexity of a time interval is considerably greater than before, as there is the monitoring requirement (conjectured to be between $\mathcal{O}(n)$ and $\mathcal{O}(n^2)$), the flight adjustment requirement ($\mathcal{O}(n^2)$), the reception requirement ($\mathcal{O}(n)$), and the handoff requirement (conjectured to be $\mathcal{O}(n^2)$).

Complexity During a Given Time of Day

We define the flux of a sector over an interval as $\Delta n = n_{\text{final}} - n_{\text{initial}}$. We interpreted an interval to be relatively short—10 to 15 min.

Entering planes have a high complexity cost, while departing planes reduce complexity. Therefore, the times of greatest complexity are not just when n is at a maximum, but rather when $d\Delta n/dt$ is at a maximum.

Potential Conflicts

The most critical aspect of the ATC's job is to reroute aircraft when a potential conflict arises. Once a trajectory is corrected, we expect that the complexity from that problem returns to 0 but the complexity of correcting the next aircraft increases, because of more limited options. We therefore conjecture that the complexity of the addition of another needed correction is related to both the total number of aircraft and the number q of previously corrected aircraft that have not departed the sector. If k is the number of aircraft still requiring course corrections, each additional potential conflict increases total complexity by $\mathcal{O}\left(aq + bk^2\right)$, with a, b constants. We expect that increasing k has a greater impact than increasing q, because of the added demand of each correction.

Effect of Software Tools

The automated tracking of aircraft would safely allow more aircraft to operate in a given sector. The motivation would primarily be economic. Air traffic would become more complex as automated software increased the ability of the system to handle a complex situation. For an ATC, the situation would be no more complex, since it could be handled with the same effort as before.

Discussion

Part A: Random Effects Model

We develop a numerical method and implement a software program to analyze simulated flight paths. The results convincingly show the strength of this approach. In all cases, the model correctly predicts what would occur.

This model has three main strengths:

- It can easily accommodate the addition of aircraft to the system.
- It allows the ATC to be confident that the sector is devoid of conflicts before aircraft even enter.
- It is general enough so that we can make refinements as new data become available for different types of aircraft and situations.

There are, however, a number of important weaknesses in this model:

- Although the theoretical approach seems straightforward and simple, actually finding the largest value across numerous three-dimensional arrays is computationally foreboding. Since this program would ideally be real-time, the solutions must be achieved quickly (under 30 s).
- The model is not based on any actual data except for FAA guidelines. Ideally, the nature of the random effects could be determined and realistic standard deviations could be used.

- We lack data as to what constitutes a dangerous probability.

Part A: The Contingency Model

This model is used when aircraft are not following specified flight plans. The model has some very strong points:

- It allows the ATC to keep track of many unplanned paths at once.
- Using an array of blocks of space facilitates the addition of multiple aircraft into the tracking system.
- Continuous updating maintains a current view of where all aircraft are projected to be.
- The model accounts for aircraft turning, when pilots are more likely to miss seeing another aircraft.

The weaknesses of this model are due to some of the constraints on its operation and some of the approximations it makes:

- Because the model approximates possible future positions with a triangle, some space is considered that the aircraft could not possibly enter. We also do not account for space that the aircraft could get to by turning less sharply.
- After a certain amount of time, the position of an aircraft along a circle begins to return to its original position, making the triangle of future positions a bad approximation. Therefore, the model is limited in how long into the future it is useful.
- The model does not account well for sharper and sharper turns.

Part B: Inherent Complexity of a Sector

The only validation possible is to ensure that the model is consistent with intuition. We expect that sectors containing more objects (airports and obstacles) are more complex, and the model supports this by suggesting that these features add to the rotationality of the sector flow. In our implementation, the magnitude of the impact a particular object (storm, mountain, airport etc.) is based purely on conjecture. With experimental data and the input of actual ATCs, the model could be calibrated.

The strength of this model is that it provides a metric for complexity that is completely general. Obviously, calibrating the model would make it more useful, but it seems unlikely that new considerations would appear that would invalidate the basic approach and the related assumptions. The weakness of this model is that although analytically sound, it is extremely difficult to implement numerically.

Part B: Aircraft-Based Complexity of a Sector

The main strength lies in providing a method for analyzing how the role of the ATC would be affected by automated tracking software. The flaws are that

- the model relies on a loose analogy between algorithmic and air traffic handling complexities,
- the model has not been calibrated, and
- the mathematics of complexity for the operations that we have defined are not well understood (that is, our operations are not numerical calculations but procedures for which the numbers of calculations are not known).

Conclusions

DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION AGENCY
AIR TRAFFIC CONTROL DIVISION
ANALYSIS SUBDIVISION
WASHINGTON, DISTRICT OF COLUMBIA

DOT-FAA 7 FEBRUARY 2000

MEMORANDUM THRU: Dr. W. Roland Hamilton, Chief of Analysis, Air Traffic Control Division

FOR: Jane Garvey, Administrator, Federal Aviation Agency

SUBJECT: Summary of report conclusions from Project Star-Chaser

1. The purpose of this memorandum is to outline the findings of Project Star-Chaser and the ramifications that for operations and personnel of the Federal Aviation Agency and Air Traffic Control System.
2. Project Star-Chaser has arrived at two computational methods for alerting ATCs to instances where two aircraft come dangerously close to one another.
   (a) Random Effects Model: This model, if applied in ATC hardware, will allow operators to determine when the flight paths of two aircraft come dangerously close. This should give ATCs confidence that, provided aircraft stick to their flight plans, there is no chance that a multi-aircraft accident can occur.
   (b) Contingency Model: This model, if implemented in ATC software/-hardware, will provide a real-time analysis of the movement of aircraft in an ATC'S sector. If an aircraft deviates from its flight plan (having to circle above an airport during severe delays, for example), the ATC has a

useful tool to help determine how the real-time behavior of aircraft lead directly to possible aircraft proximity conflict.

3. Perhaps more important than developing these computational tools, however, Project Star-Chaser has been largely focused on the systemic complexity associated with Air Traffic Control. We have broken the analysis of this complexity into several portions.

   (a) Inherent Complexity of a Sector: By examining the impact that objects have on a sector of airspace (namely, airports and obstacles), we have developed a means to analyze the complexity of a sector. This is the key to evaluating our current sector arrangement and determining if geographical boundaries should necessarily determine how the National Airspace is subdivided.

   (b) Instantaneous, Time Interval, and Time of Day Complexity: These three areas are iterative refinements of how the aircraft in the sector lead to increased complexity. Even with employing a minimum spanning tree as a best case scenario, the complexity of an ATC's workload is still $\mathcal{O}\left(n^2\right)$. This means that, under the present circumstance, increasing the number of aircraft leads to a much higher complexity for the ATC.

   (c) Impact of Collision Corrections: The Project further examined the complexity model to account for how complexity is affected by the number of course corrections needed. We find that this complexity is $\mathcal{O}\left(q+k^2\right)$, where q is the number of corrections already made and k is the number of corrections outstanding. This result suggests that a backlog of corrections greatly increases the complexity of the ATC's job.

   (d) Impact of Advanced Information Systems: We predict that the use of more advanced and autonomous hardware/software packages will reduce the job complexity for the ATC.

4. Advances will reduce the workload placed on ATCs and will increase efficiency and the volume of air travel.

   Unfortunately for ATCs, the end result of system improvements is no change. Though advances in guidance, tracking, and other technological areas seem to offer hope to reduce the stress and worry of the job, market and other economic equilibrium forces will quickly return the system to its maximum safe carrying capacity.

5. The Point of Contact for this memorandum is the undersigned at `J_Gibbs@Star_Chase.org`.

JOE W. GIBBS, Ph.D.
Head Functional Analyst,
Project Star-Chaser

Appendix A: Identification of σs

We assume that the net effect of variations in an aircraft's actual position compared to its predicted position is zero, since they are nonsystemic. In addition, the additive combination of random variations strongly suggests that the error term is normally distributed.

We do not have a good way of measuring the standard deviations in the three coordinate directions. Instead, we "reverse engineer" the standard deviations based on FFA guidelines for the minimum separation of aircraft in flight.

Table A1.

FAA guidelines and corresponding standard deviations.

Orientation	Altitude	Minimum separation	Standard deviation (km)
Laterally	N/A	5 mi	1.563
Vertically	<29,000 ft	1,000 ft	0.304
Vertically	>29,000 ft	2,000 ft	0.430

We determine the values for the standard deviations in **Table A1** by assuming that the FAA set guidelines so that there would be a .999 chance that the actual position of the plane would be within the accepted range; hence we use for each dimension the equation

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\text{min dist}}^{\text{min dist}} \exp\left(\frac{-s^2}{2\sigma^2}\right) ds = 1 - .999$$

and solve for σ using MathCad.

Appendix B: Results of Simulations

Figure B1 (next page) shows the result for the collision probabilities procedure on data for planes whose paths intersect at the same point at the same time. **Figure B2** shows the result for a near-miss, paths that intersect in space but not at the same time. The maximum probability in the near-miss case differs from the maximum probability in the collision case by 11 orders of magnitude.

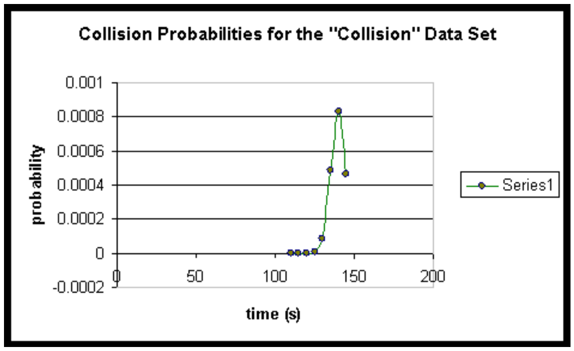


Figure B1. Collision probabilities for the “collision” data set.

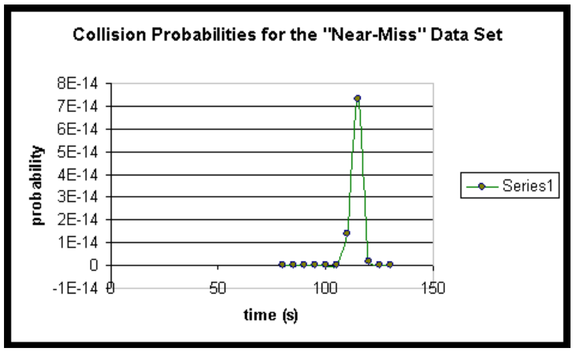


Figure B2. Collision probabilities for the “near-miss” data set.

References

Baird, D.C. 1995. *Experimentation: An Introduction to Measurement Theory and Experimental Design*. 3rd ed. Englewood, NJ: Prentice Hall.

Denver Air Route Traffic Control Center (ARTCC) Homepage. 2000. `http://www.tc.faa.gov/ZDV/` .

Devore, Jay L. 1995. *Probability and Statistics for Engineering and the Sciences*. 4th ed. Boston, MA: Duxbury Press.

Eppstein, David, and Jeff Erickson. 1999. Raising roofs, crashing cycles, and playing pool: Applications of a data structure for finding pairwise interactions. Extended abstract at *Proceedings of the 14th Annual ACM Symposium on Computational Geometry* (1998), 58–67. *Discrete and Computational Geometry* 22 (4): 569–592. `\http://www.uiuc.edu/ph/www/jeffe/pubs/cycles.html` .

Hibbeler, R.C. 1998. *Engineering Mechanics: Dynamics*. 8th ed. Upper Saddle River, NJ: Prentice Hall.

Horowitz, Ellis, Sartaj Sahni, and Sanguthevar Rajasekaran. 1997. *Computer Algorithms / C++*. New York: Computer Science Press.

Rosen, Kenneth H. 1999. *Discrete Mathematics and Its Applications*. Boston, MA: WCB McGraw-Hill.

Schey, H.M. 1997. *Div, Grad, Curl, and All That: An Informal Text on Vector Calculus*. 3rd ed. New York: W.W. Norton.

Spiegel, Murray, and John Liu. 1999. *Mathematical Handbook of Formulas and Tables*. 2nd ed. New York: McGraw-Hill.

Stewart, James. 1998. *Calculus II. Special Edition for the United States Military Academy, Fall/Spring 1998/1999*. Boston, MA: Brooks/Cole Publishing Company.

Ugural, Ansel, and Saul Fenster. 1995. *Advanced Strength and Applied Elasticity*. 3rd ed. Upper Saddle River, NJ: Prentice Hall.

You Make the Call: Feasibility of Computerized Aircraft Control

Richard D. Younger
Martin B. Linck
William P. Woesner
University of Colorado
Boulder, CO

Advisor: Anne M. Dougherty

Introduction

We investigate whether some of the work done by air traffic controllers (ATCs) could be handled by computers. Automated systems could act as watchdogs, heading off crises before they become catastrophes. Specifically, we investigate at what point an ATC must take charge of a situation to avoid catastrophe, what sort of decision must be made to remedy the situation, and how much stress is involved.

Objectives

- Define a minimum safe distance between aircraft.
- Develop a numerical model of air traffic around a busy airport.
- Assess system complexity and corresponding ATC workload under a variety of circumstances.
- Develop aircraft guidance algorithms that minimize controller stress.

The System

An ATC has three main tasks as an aircraft approaches an airport, all of which must be carried out as quickly as possible [Wood 1983]:

The UMAP Journal 21 (3) (2000) 285–300.

- Ensure that the aircraft does not collide with another aircraft or any other obstacle.
- Ensure that the approaching craft is inserted smoothly into the traffic around the airport, with a minimum of disruption to the flight paths of other aircraft.
- Guide the aircraft onto a runway, again with a minimum of disruption to the rest of the traffic around the airport.

Special cases, such as aircraft experiencing mechanical malfunctions, medical crises, or fuel shortages, must be dealt with, and changing weather conditions must be taken into account.

To make the simulation concrete, we model Denver International Airport (DIA); our methods could be extended to nearly any air traffic control center.

Each ATC is assigned a specific type of task. Thus, one set of controllers assigns flight paths to incoming aircraft, another guides those aircraft to holding patterns or landing approaches, and yet another guides planes to a safe landing. As an aircraft passes from one controller to another, the pilots switch radio frequencies. Each frequency belongs to a specific controller, and each ATC watches a radar screen on which icons represent flights for which that ATC is responsible. The tower controllers, who are in charge of landings, can see the aircraft and so are not as dependent on radar information [FAA 1999].

At DIA, it is common practice to route all incoming flights on north-south and east-west vectors, since prevailing wind conditions usually favor these approaches [Wood 1983]. Departing flights must use the same runways as incoming flights but once airborne are routed out of DIA airspace on northeast-southeast and northwest-southwest vectors, to minimize conflicts between incoming and outbound aircraft. **Figure 1** shows a diagram of the general approach and departure vectors. To make its final approach and land on a runway, each aircraft has to pass a point in space approximately 5 mi from the end of the landing runway and approximately 2 mi (10,000 ft) above ground level.

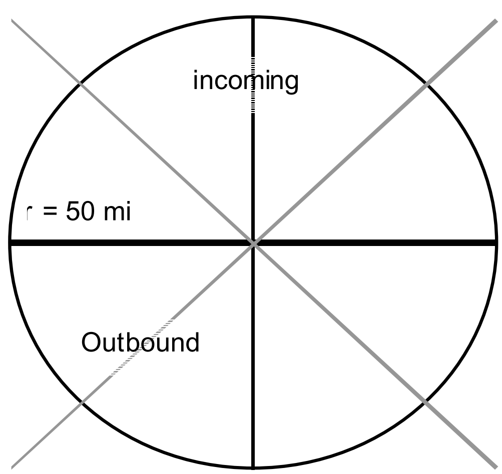


Figure 1. Airspace approach and departure diagram.

Assessing Safety

We use real data wherever possible; we consult the Federal Aviation Regulations (FAR) [FAA 1999] whenever technical questions arose.

Federally regulated Instrument Flight Rules (IFR) state that the minimum safe distance between adjacent aircraft is 1,000 vertical feet and 3 mi of horizontal distance, when aircraft are moving at landing speeds in close proximity to each other and to the airport. Since the runways at DIA are 4,330 ft apart, this minimum safe distance must be ignored on final approach and on runways.

Assessing Complexity, ATC Workload and Stress

Goode and Machol [1957], writing about large-scale queueing systems, describe complexity as "the extent to which any given attribute of a system will affect all the others if it is changed." In a complex system, all the variables are tightly linked; one could not, for example, change the position of an aircraft without immediately having to change some characteristic of most of the other aircraft in the system. Unfortunately, measuring this type of complexity is difficult, since there is no obvious set of measurable system variables for assessing how closely each variable depends on all the others.

Further, we are interested not merely in the complexity of the system but in the stress and fatigue that the system is likely to cause its ATCs. Surprisingly, there is no accepted set of factors that cause ATC stress, fatigue, and error. Some researchers, such as Redding [1992], have concluded that the number of incidents was highest when ATC workload was actually moderate to intermediate. On the other hand, Morrison and Wright [1989], reviewing NASA data, report that ATCs make more mistakes when the workload is at its highest. One innovative study of the phenomenon of ATC fatigue is that by Brookings et al. [1996]. Their test subjects were Air Force ATCs who were asked to play an air traffic control simulation and were exposed to scenarios of varying difficulty. One scenario was an "overload scenario," in which they were asked to coordinate the movements of 15 aircraft at once. As the ATCs attempted to deal with each scenario, their heart rates, blink rates, and brain activity were monitored. Brookings et al. concluded that there was a correlation between workload and operator stress and that the likelihood of error—in the form of separation errors (not enough room between planes), fumbled approaches, and botched handoffs—was directly linked to ATC stress. As a result, we assume for our simulation that ATC stress should be minimized and that there is no "optimal stress level" [Redding 1992] at which ATCs should operate.

Based on our literature survey, we conclude that there are four measurable factors that influence ATC stress levels:

- n, the number of total planes in the airspace around the airport.
- f, the number of separation errors currently occurring in the airspace around the airport. We assume that a single separation error causes as much stress

as 20 extra aircraft in the airspace.

- d, the *smallest* distance between planes currently on the map.
- a, the *average* distance between aircraft. If all the aircraft are well distributed throughout the airspace of the airport, the ATCs are likely to experience less stress than they would if they were all clustered together.

Our formula for the stress-causing complexity experienced by ATCs is

$$C = n + 20f + 100/d + 100/a.$$

Queueing Theory, Stochastic Input, and Algorithm Design

The airport is a queueing system. A queueing system has servers that handle input, perform some function on the input, and then pass the input to some other part of the system. Input is not created or destroyed in the system. The servers are usually referred to as *channels*. The five active runways of DIA are the channels of the system.

When the number of servers is inadequate to the amount of input they are called upon to handle, a queue develops. In the case of an airport, the holding patterns in which aircraft wait for clearance to land are the queues of the system.

When the amount of traffic is stochastic (determined by a probabilistic distribution) and the input is discrete, the input density is often described by a Poisson distribution. The amount of time for each channel to process an input need not be constant. The standard numerical approach to modeling would involve setting up an input generator, which would send us airplanes according to a density function. We would then set up our channels, and the amount of time to process each plane would vary, probably according to a normal distribution. We would also set up holding patterns, to which our planes could be sent when there are no runways available. We could then let the program run and see how queues develop as time passes.

Unfortunately, this simple approach doesn't allow us to investigate all the questions posed by the problem statement. The objective of the simulation should be to develop and test algorithms to monitor the position and speed of aircraft and to alter flight paths to minimize the workload and stress of ATCs. If we treat airplanes simply as inputs, without them becoming objects maneuvering in space, we cannot adequately test those algorithms. So, a more ambitious approach is called for, one that allows us to “create airplanes” stochastically, maneuver them, send them to holding patterns, land them, and assess the value of stress-causing complexity as the simulation runs.

The Model

An airport is a continuous system; each airplane continuously changes position and velocity. However, there are discrete events that characterize the system, such as takeoffs, landings, and handoffs. Both continuous and discrete characteristics of the system can be described by a discrete model provided the time resolution is fine enough. Recognizing this, we develop three numerical simulations of aircraft behavior, which test:

- the minimum-safe-distance assumption,
- guidance algorithms on aircraft landing, and
- guidance algorithms for aircraft entering and maintaining a holding pattern.

Minimum-Safe-Distance Simulation

Consider two aircraft traveling on parallel flight paths at an airspeed of 300 mph, well below the cruising speed of commercial airliners. Let the vertical axis be the z-axis, the axis parallel to the flight path be the x-axis, and the axis perpendicular to both be the y-axis. Wind and other factors perturb the velocity vectors of the planes according to (we assume) a normal distribution with mean zero. A large standard deviation might correspond to the planes flying in heavy weather, a small one to calm weather.

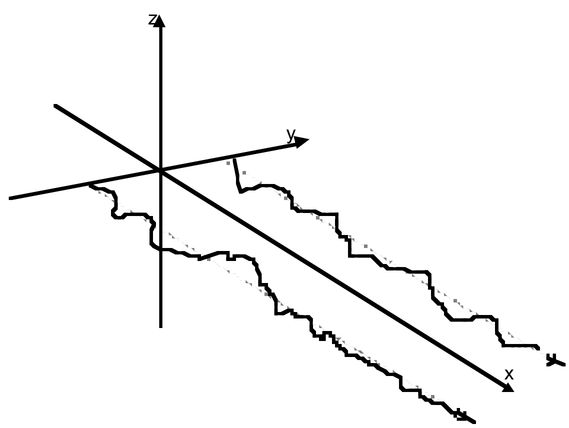


Figure 2. Aircraft separation simulation.

If we did not correct the aircraft's course, the plane would diverge from its flight path and move in a random walk as time passes. Our program determines the distance between the aircraft and its flight path and changes the velocity vector to bring the plane back on course. Airplanes have maximum rates of acceleration in any direction, and we assume that the aircraft can change velocity in any direction by no more than 10 mph/s. The result is a corrected random walk, with the step size normally distributed and a finite correction to each step.

Since the wingspan of an airliner is approximately 200 ft and turbulence effects surround the aircraft, we assume that if the airliners come within 250 ft of one another, they collide. We assume for the sake of simplicity that this holds true in the vertical direction as well. To test our minimum-safe-distance assumption, we fly planes next to each other for 100 h; if the likelihood of collision is less than 0.05%, we consider the distance safe for the given weather conditions.

Developing Aircraft Guidance Algorithms

A flow diagram for our airport is shown in **Figure 3**.

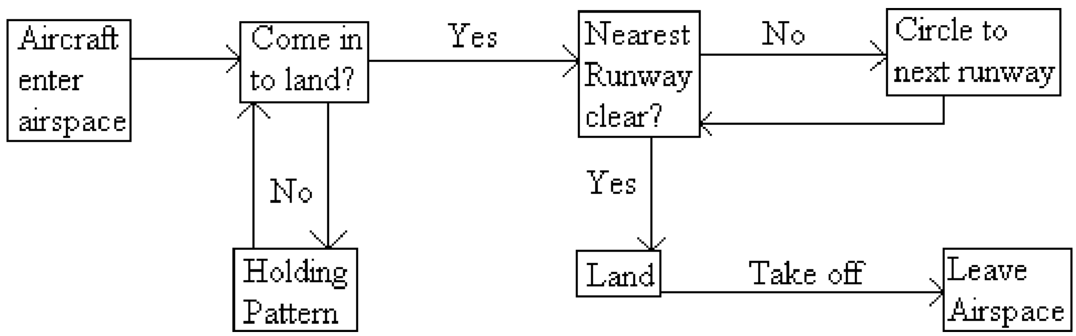


Figure 3. Flow diagram for an airport.

Airplanes enter the airspace according to a stochastic distribution. If a landing approach is free, they proceed towards it. If no runway is clear or if the airspace is too crowded, they are sent to a holding pattern (our queue). The amount of time in the holding pattern depends on the rates at which planes land and planes enter the airspace.

We assume that the scheme used to select the next aircraft cleared for approach to landing is FIFO (first in, first out). If we wished to take into consideration factors such as fuel or other flight emergencies, each plane would have to be assigned a priority and planes would be pulled from the queue according to priority.

Once an aircraft has been cleared for approach, it must select a runway. If the nearest runway is not free, the aircraft must circle until one is. Once it finds a clear runway, it may land. The runway is then occupied for some (perhaps stochastic) amount of time. The aircraft then spends some (perhaps stochastic) amount of time on the ground and takes off again.

The process can be divided into boxes, or areas. How much stress airplanes in a given area cause to the ATCs depends on the number of airplanes and the extent to which guidance algorithms manage to control the aircraft. Aircraft just entering or just leaving the airspace are unstressful, since they are presumably spaced out along a very large circumference. Aircraft within 5 mi of a runway that are coming in to land are admitted only when a runway is clear; they are handled by ATCs who monitor those flights visually as well as via radar. The consequences of a mistake in this function are so high that we can safely assume that humans will handle this task for the foreseeable future.

There remain two areas in which our algorithms might ease controller stress:

- aircraft that have received clearance and must line up for a landing. **Figure 4** shows runway checkpoints, through one of which an aircraft must pass to land.

- aircraft that have not been cleared to land and must be sent to a queue. **Figure 5** shows the set of checkpoints that comprise the holding pattern. An aircraft can enter the holding pattern at any checkpoint, and, if properly guided, proceeds to the next checkpoint in the sequence. It cycles through the sequence until given clearance to approach. All aircraft move through the sequence in the same direction, to avoid head-on collisions.

We design algorithms to maintain aircraft spacing while guiding those aircraft, either toward a runway checkpoint, or through the circular set of checkpoints that form the queue.

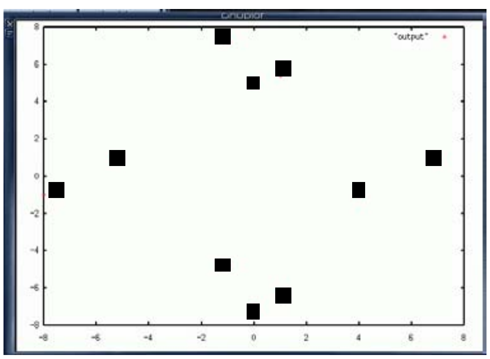


Figure 4. Runway checkpoints.

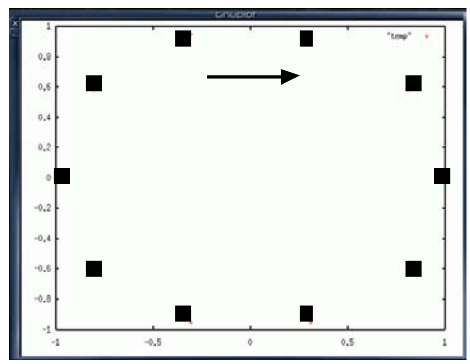


Figure 5. Queue checkpoints. Arrow indicates flight direction.

The Algorithms

Single-Avoidance

This algorithm determines the distance between each aircraft and that aircraft's next checkpoint and orients the aircraft toward the checkpoint. It also evaluates the distance between that aircraft all other aircraft; if any distance is equal to or smaller than the minimum distance (3 mi horizontal, 1,000 ft vertical), it then orients the aircraft directly away from the dangerously close airplane (the aircraft in the airspace longer changes course) without changing speed. Once the distance again exceeds the minimum safe distance, the aircraft that had to change its course looks for its nearest objective (which need not be the same objective that it was originally approaching) and changes course toward that objective. An example of aircraft being guided toward a landing checkpoint by the single-avoidance algorithm is shown in **Figure 6**.

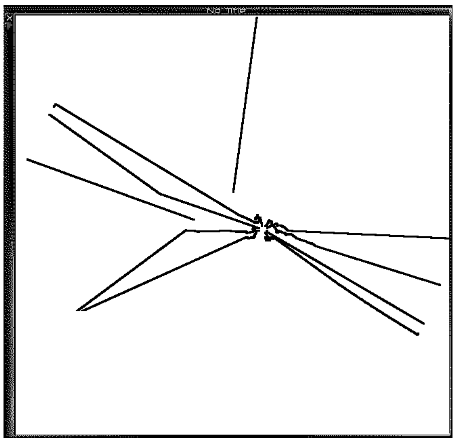


Figure 6. Aircraft converging on runway checkpoints while guided by the Single-Avoidance Algorithm.

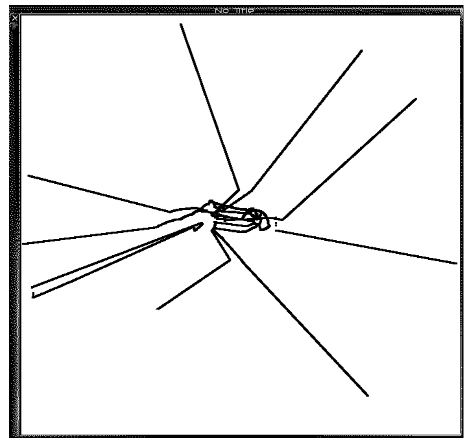


Figure 7. Aircraft moving toward landing checkpoints under the Double-Avoidance Algorithm.

Double Avoidance

If the distance between two aircraft drops below the minimum safe distance, *both* aircraft head away from each other, changing courses equally and in opposite directions without changing speeds. **Figure 7** shows an aircraft being guided toward a landing checkpoint by the Double-Avoidance Algorithm.

Single-Vector Repulsion

Unlike the first two algorithms, this algorithm constantly measures the distance between the aircraft under its control and that aircraft's nearest neighbor. It alters the course of that aircraft before a separation error can occur. The scheme used to correct the aircraft's flight path is as follows.

In each time step, the algorithm determines the direction in which the aircraft must fly to reach the nearest checkpoint. It thus establishes a velocity vector $\vec{V}$ for the aircraft, whose magnitude cannot change but whose direction is corrected in every time step. It then finds the vector $\vec{D}$ that connects the craft under guidance with its nearest neighbor and calculates a correction vector $\vec{A}$ whose direction is the same as that of $\vec{D}$ but whose magnitude is

$$||\vec{A}|| = \frac{b}{||\vec{D}||^2},$$

where b is a scaling constant. If b is large, the correction is severe, even at large distances; if b is very small, we run the risk that flight paths will not be adjusted quickly enough and the aircraft will come to close to each other.

The velocity vector $\vec{V}$ is corrected to $\vec{V_c} = \vec{V} - \vec{A}$. The result is that every aircraft is repelled by the aircraft nearest to it, with increasing intensity as the distance between the adjacent aircraft becomes smaller. At the same time, the aircraft remains attracted to its objective.

If there are only two aircraft in the simulation, and both are headed for the same objective, they behave like a pair of negatively charged ions approaching a large positively charged sphere (this analogy breaks down when there are more than two aircraft). **Figure 8** shows aircraft maneuvering toward an objective while guided by the Single-Vector Repulsion Algorithm. Note that most of the aircraft maintain much better spacing than with the previous algorithms.

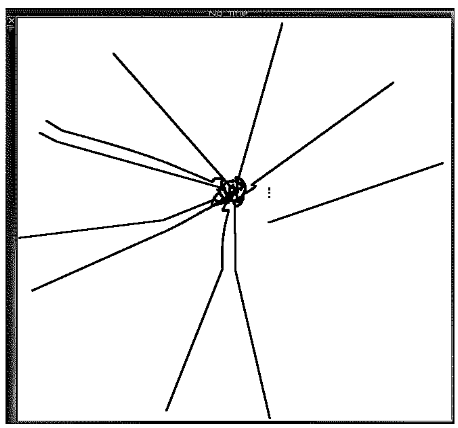


Figure 8. Aircraft maneuvering toward landing checkpoints under the guidance of the Single-Vector Repulsion Algorithm.

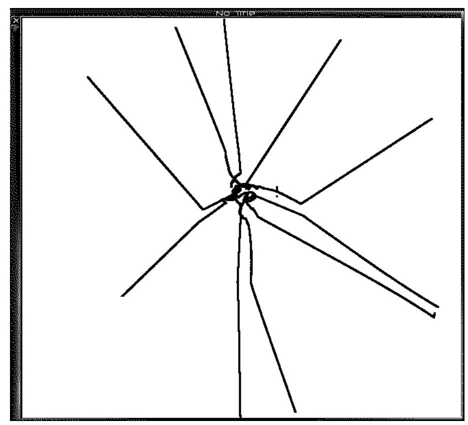


Figure 9. The Multiple-Vector Repulsion Algorithm.

Multiple-Vector Repulsion

The Multiple-Vector Repulsion Algorithm has the same vector mechanics as the Single-Vector Repulsion Algorithm, but each airplane is repelled not just by its nearest neighbor but by every other airplane in the airspace. The behavior of a number of aircraft headed toward a single objective would be roughly analogous to a group of negatively charged ions heading toward a large positively charged sphere. With more than one objective, however, the analogy is less apt, since each "ion" is attracted only to the nearest sphere. **Figure 9** shows 10 aircraft approaching a checkpoint under the guidance of the Multiple-Vector Repulsion Algorithm.

Testing the Algorithms

The Landing Approach Test

In this test, aircraft enter the airspace according to a normal distribution with mean 120 s and a standard deviation of 60 s. The points on the circumference of the airspace are evenly and randomly distributed. The airspace is centered on the airport and has a radius 50 mi. Each aircraft enters the airspace at an altitude of 6 mi (close to cruising altitude) and descends to one of 10 runway checkpoints (each of the 5 runways can be approached from either end), each located 5 mi from the end of a runway and at an elevation of 2 mi. Once an airplane reaches a runway checkpoint, its velocity instantly becomes zero. This ensures that no further aircraft can reach that checkpoint while the plane remains in place. After a set period of time (1 min in our simulation, based on the FAR [FAA 1999]), the airplane disappears and can be described as having landed. The runway is then clear and a new aircraft can occupy that checkpoint. The value for ATC stress is calculated at each time step.

The Queueing Test

Aircraft enter the airspace just as they do in the landing approach test. They descend to a set of queueing checkpoints and maneuver around those checkpoints in a holding pattern. Each holding pattern has a saturation level of airplanes, equal to the perimeter of the pattern divided by twice the minimum allowable distance between aircraft. Aircraft are added to the holding pattern stochastically until saturation is reached. We calculate the amount of ATC stress that the holding pattern generates. **Figure 10** shows 10 aircraft descending toward and entering the holding pattern under the control of the Single-Vector Repulsion Algorithm.

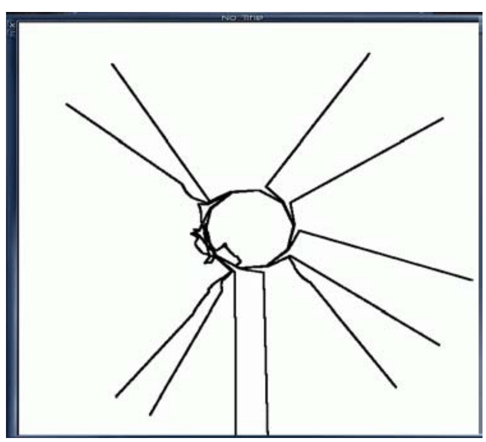


Figure 10. Aircraft being maneuvered in to a holding pattern by the Single-Vector Repulsion Algorithm.

Simplifying Assumptions in Testing the Algorithms

- All aircraft behave like commercial air carrier aircraft.
- The velocity of all aircraft is 300 mph.
- The minimum vertical separation between aircraft is 1,000 ft.
- The minimum horizontal separation between aircraft is 3 mi.
- The turning radius of all aircraft is negligibly small compared to the overall airspace.
- An aircraft has reached a checkpoint when it has passed within 1 mi of that checkpoint.
- Weather conditions do not affect the behavior of aircraft and are not considered.
- No aircraft is given any special priority over any other; fuel and emergency considerations are therefore ignored.
- There is no coordinating intelligence at work. The human ATCs can watch the simulation (and be stressed by it), but all actions of the aircraft are determined by the algorithm being tested.
- Aircraft are incapable of acting without instructions from the control algorithms. In essence, the pilot of each aircraft blindly and unquestioningly follows the orders given by the algorithm.

The following assumptions apply only to the Landing Approach test for all algorithms:

- A runway approach checkpoint remains occupied for 1 min.
- Any aircraft that reaches an approach checkpoint lands; there are no failed landing attempts.
- Once an aircraft has landed, it ceases to interact with any other aircraft and disappears from the simulation.
- Outbound aircraft do not interact with inbound aircraft and hence do not appear in the simulation. **Figure 11** shows that with our current checkpoint system, corridors develop in which there is no inbound traffic. So we assume that outbound traffic passes through these corridors and need not be accounted for.

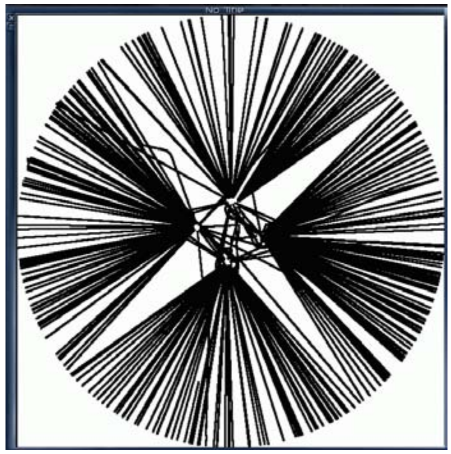


Figure 11. Typical traffic pattern when aircraft head toward the checkpoints from the periphery. Note the open corridors, along which outbound traffic can be routed.

Results and Commentary

Minimum-Safe Distance Simulation

Figure 2 shows the flight paths of two aircraft as they travel next to each other for 5 min, together with ideal flight paths that are separated by 3 mi. The results of many such simulations are given in **Figure 12**. Each data point represents the average of 150 runs, with each run lasting 100 h. The components of each aircraft's velocity are each disturbed by components with a mean of zero and with standard deviations as noted in the legend. The disturbances correspond to moderate, heavy, and severe weather conditions.

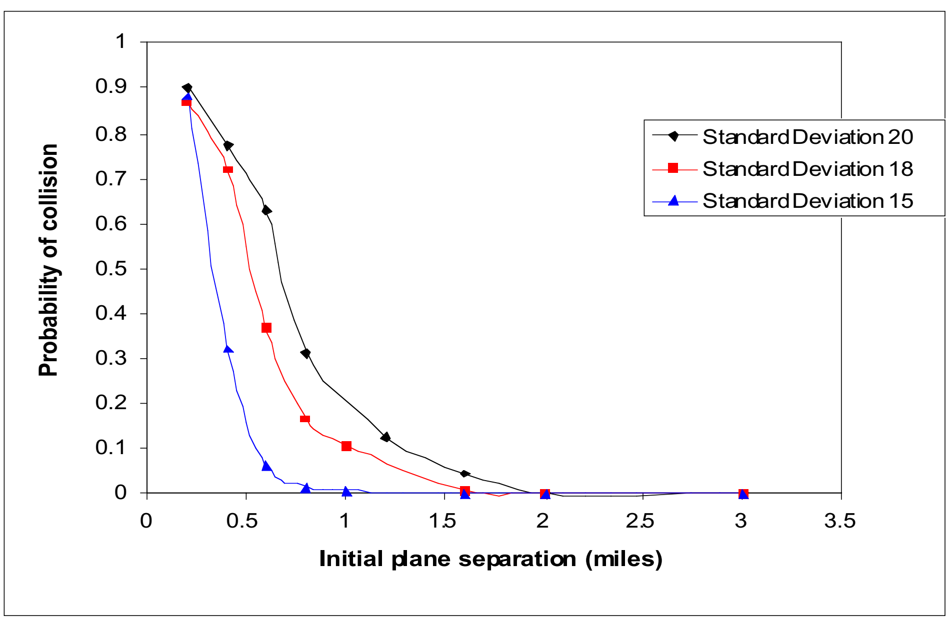


Figure 12. Probability of collision in a 100 h run as a function of initial separation.

The likelihood of collision is well under our 0.05% criterion when the separation distance is 3 mi. In fact, in the course of all 450 runs of 100 h, over all three disturbance levels, no collisions occur. The likelihood of collision increases exponentially as the horizontal separation distance decreases and becomes appreciable at 1.8 mi, where the first collision occurs. We conclude that FAA regulations give a secure safety margin and so abide by them for the rest of the simulation.

Holding Pattern and Landing Approach Simulations

There are no statistically significant differences among the stress levels for our four test algorithms, neither for holding pattern nor landing approach simulations. The holding pattern simulation achieves equilibrium with little to no stress and then the stress levels out. Since the landing simulation is much more complicated, one would expect sharp spikes and peaks in the stress, and this is what we see. **Figures 13** (generated using the Single-Vector Repulsion Algorithm) and **14** (generated using the Multiple-Vector Repulsion Algorithm) show sample graphs of stress vs. time for the queueing and landing simulations, respectively.

In **Figure 13**, the stress rises as planes are added; when the pattern is saturated, the stress level becomes steady. There are few collision warnings; the rise in stress is caused by the decreasing proximity of the planes as they approach the holding pattern.

Figure 14, however, is fascinating. For $0 < t < 5{,}000$, the graph is piecewise continuous, but the rest is totally discontinuous. In our definition of stress, there are two continuous terms and two discrete terms. Over the first 5,000 s of the simulation, the continuous terms play an important role in stress; but during the latter 5,000 s, the continuous terms die off, leaving the discrete terms to determine the stress.

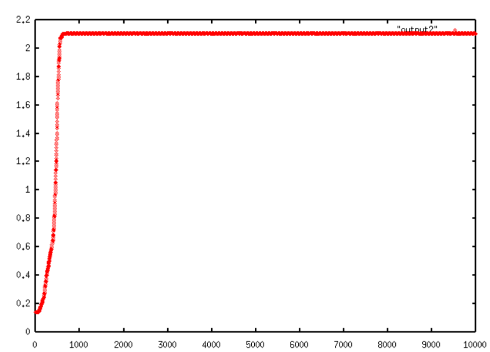


Figure 13. Stress vs. time for a queueing simulation.

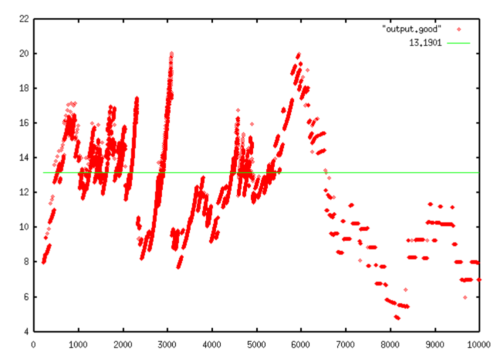


Figure 14. Stress vs. time for a landing simulation.

Table 1 gives summary statistics for both simulations run with each of the controlling algorithms. Each scenario was repeated 100 times. The seed for the pseudorandom number generator was held constant across the controlling algorithms, so that each algorithm received the same input.

Table 1.
Descriptive statistics from the simulations.

	Landing Simulation		Queueing Simulation	
	Mean	SD	Mean	SD
Single Avoidance	12.8	0.8	4.1	1.9
Double Avoidance	14.2	1.7	4.1	1.9
Single-Vector Repulsion	14.2	1.7	4.1	1.9
Multiple-Vector Repulsion	13.2	1.6	4.8	2.2

The most interesting result is that the Single Avoidance Algorithm performs better than the other more clever algorithms. Here we can draw a parallel to the phenomenon from computer science known as *deadlock*. Deadlock occurs when two or more processes cannot continue executing because each process is requesting resources owned by the other process [Nutt 1999]. Despite much research and many clever algorithms, the standard way to handle deadlock in a computing environment is to just pick a winner, usually the process that has been waiting the longest. The Single Avoidance Algorithm is analogous: When two planes request the same airspace, only one can be awarded it; the algorithm picks a winner and vectors the loser away.

To make the Multiple-Vector Repulsion Algorithm more efficient, we increase the repulsion factor between planes; if planes are kept farther apart, they will not incur collision warnings. The idea is good but the results are not encouraging. **Figure 15** is a plot of stress vs. time for the landing simulation using the Multiple-Vector Repulsion Algorithm with increased repulsion.

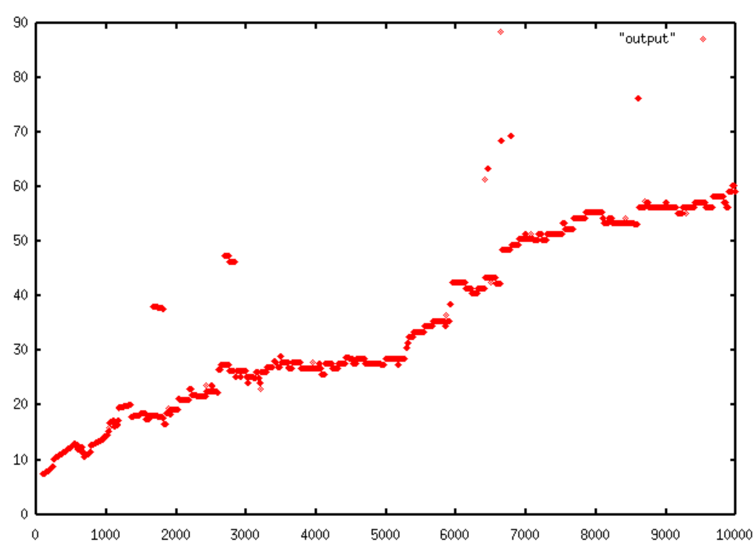


Figure 15. Stress vs. time for the landing simulation using the Multiple-Vector Repulsion Algorithm with increased repulsion.

We see a monotonically increasing stress level—certainly not the desired result. We can explain this phenomenon too in terms of deadlock. Imagine that two airplanes approach a checkpoint on opposite approach vectors. When they get close enough to each other, the repulsion factor makes them turn around. When they again get far enough away from each other, they turn around and fly back toward the checkpoint; and the cycle repeats. The planes become deadlocked, so very few planes can land. Hence, the number of planes in the airspace increases, leading to increasing stress.

The queueing simulations generate very little stress. Hence, software should be able to take control of a plane, vector it into a holding pattern, and keep it there.

Strengths and Weaknesses

More factors contribute to this system than we are able to consider in a weekend. However, our model is modularized to the point where modules can be run independently of each other, making it easy to focus on specific parts of the model.

It took our workstation approximately two hours to generate the data for our limited model, but parallel processors could each handle a set of planes.

Our model covers all of the major elements of an airport simulation; while it is based on DIA, it can easily be generalized.

Conclusions and Recommendations

We conclude from our simulation that *FAA guidelines for aircraft separation give a generous margin* for navigational, technical, and pilot error. Perhaps they could be relaxed, especially for well-controlled parallel landing situations in good to moderate weather.

The air traffic control problem exhibits many traits common to NP-complete problems:

- There is no deterministic way to pick the best solution.
- Humans appear to control air traffic much more effectively than computers could.
- Changes in the control algorithm for our simulation do not seem to impact the quality of the solution at all.

These features lead us to speculate that this problem is indeed NP-complete.

The computer does, however, handle the queueing problem fairly well; planes are vectored into the queue and held there with minimal stress. These results are promising; they suggest that *though computers should not replace humans as the primary means of air traffic control, computers might be capable of handling the more mundane tasks*.

References

Airports Data Base — Denver International Airport. 2000. `http://www.boeing.com/assocproducts/noise/denver.html` .

Brookings, J.B., G.F. Wilson, and C.R. Swain, 1996. Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology* 42: 361–77.

Federal Aviation Administration. 1999. *Federal Aviation Regulations*. Washington, DC: Government Printing Office.

Goode, Harry H., and Robert E. Machol. 1957. *System Engineering*. New York: McGraw-Hill.

Morrison, R., and R.H. Wright. 1989. ATC control and communication problems: An overview of recent ASRS data. In *Proceedings of the Fifth International Symposium on Aviation Psychology*, 902–907.

Nutt, Gary. 1999. *Operating Systems: A Modern Perspective*. Reading, MA: Addison-Wesley.

Redding, R.E. 1992. Analysis of operational errors and workload in air traffic control. In *Proceedings of the 36th Annual Meeting of the Human Factors Society*, 1321–1325.

Wood, Kenneth Ray. 1983. Airport activity simulation. Master of Science thesis, University of Colorado. Boulder, CO.

Judge's Commentary: The Outstanding Air Traffic Control Papers

Patrick J. Driscoll
Department of Mathematical Sciences
U.S. Military Academy
West Point, NY 10996
ap5543@exmail.usma.army.mil

Introduction

The Federal Aviation Administration (FAA) Air Traffic Control problem was an interesting mix of quantitative and qualitative inquiry. Teams used a variety of modeling approaches to resolve the qualitative issue of how to model the complexity of an air traffic controller's job; they apparently discovered that this type of question is often the most challenging, the most interesting, and frankly, the most fun to tackle.

The Approaches of the Top Papers

The top papers in this year's contest did so with a flair of creativity and obvious thoughtful consideration for all dimensions of the modeling process. Many recognized that the number of aircraft in a controller's sector must ultimately be a significant contributing factor to this complexity. A good number of papers further refined this notion to include dimensions such as workload, relative proximity of aircraft, and number of aircraft flight path adjustments required per unit time, among others.

Many papers recognized that aircraft conflict occurs in pairs and proceeded to assess the maximum number of possible conflicts for a given scenario. Some papers chose to divide the overall airspace into vertically separated layers and then developed conflict algorithms for the 2-D problem on a particular layer as opposed to using a 3-D model from the start. However, several papers fell

The UMAP Journal 21 (3) (2000) 301–304.

back on a tacit 3-D model for complexity without realizing that their earlier results did not extend to this case. The most common approach to address the 3-D problem directly was to create an inner collision space (either a sphere or rectangle) around the aircraft and then use a larger alert space (containing this inner collision space) for early warning. The difference in radii between the two spaces was used as a measure for air traffic control conflict reaction time, so that a controller could adjust one of the aircrafts' courses without causing excessive internal forces on the aircraft, its passengers, or its cargo.

There was a wide range of techniques used by teams to represent and identify potential conflicts along aircraft flight paths. Some reduced this problem to a time-parameterized vector-intersection problem in the 2-D plane, and others did the same for the 3-D case as well.

There were several more extensive approaches worth noting:

- One paper assumed that a drift error exists from wind effects, weather, and turbulence along an aircraft flight vector in three-space and applied a probability distribution to this error. This enabled the team to construct a stochastic simulation to test their model design.

- Still another paper incorporated both straight and curved parametric flight paths into their methodology. In both cases, if the flight paths of two aircraft drifted sufficiently close to cause an alert space violation, a controller was presumed to take corrective action to prevent intrusion into the collision space. The number of these alerts occurring for a given aircraft density in traffic then became a component of complexity measurement.

Other Papers

The papers that did not rise to the top possessed glaring omissions or oversights that typically manifest themselves when teams run out of time, fail to properly identify all of the questions being asked, or develop complex mathematical representations and then find themselves lacking the ability to solve them. There seemed to also be a bit of uncertainty in some papers as to what exactly it means to *analyze* a model.

Teams that dismissed portions of the problem based on the claim that "FAA ATC conflict software already exists" missed the point: The FAA knows what they have in their repertoire of tools; they are looking for other ideas and approaches that might facilitate a better solution than they currently have. If the head of the FAA were completely satisfied with the status quo, there would be no problem to be solved.

The Need for Verification

The Model Must Produce Results...

Certain fundamental elements of modeling are consistently present in quality applications of the problem-solving process. Without these, technical reports have noticeable gaps in the information that they provide—gaps that call into question the validity, voracity, credibility, and applicability of the results being presented. This situation is similar to courtroom testimony; judges and juries have a difficult time believing a witness unless sufficient evidence is presented to support the witness's testimony. Consequently, if an MCM team's principal effort is to construct computer code to simulate an air traffic scenario, *they must present results that provide evidence that their code/model actually ran and yielded the information sought*. Analyzing the output of a model provides a basis for determining if the modeling approach chosen was reasonable.

...Which Must Be Presented and Analyzed...

Simply put, after creating an acceptable mathematical representation (system of equations, simulation, differential equations, etc.) of a real-world event, this representation (model) must be tested to verify that the information it produces (solutions, simulation output, graphics, etc.) makes sense in the context of the questions being asked and the assumptions made to create the mathematical representation. *It is insufficient to present such a representation without this additional evidence*. Once a mathematical model is created, symbolic, graphical, and/or numerical methods must be used to produce evidence that the model works. Many of the best papers did so using a combination of these three approaches; some teams wrote C++ code or used spreadsheets, while others used a computer algebra system such as MAPLE or MathCad as their workbench.

...and Compared with Clearly Stated Assumptions

Far and away, papers reaching the final round of judging paid a good deal of attention to stating their assumptions clearly, explaining the impact of each assumption and why they felt it was necessary to include it in their model development. They were also very careful not to assume away the challenging and information-relevant portions of the problem posed. Teams' increased sensitivity to this aspect of the modeling process has consistently improved over the years to the point that it is a hallmark of most modeling efforts today. From a judging perspective, it is far easier to follow the logical construction of these teams' models and to identify what they were attempting to do. However, even a few of the best papers mistakenly placed key pieces of information in appendices rather than in the section of the paper where supporting evidence was desperately needed.

Use of Existing Research

Teams are increasingly adept at using the Internet to find credible, reliable information sources to support their modeling efforts. There is a good deal of room for improvement in team papers as to how best to incorporate this information properly into a technical report, especially for a team that perceives that it has struck the motherlode of reference sources. Incorporating others' work without diluting one's own effort is challenging. However, parroting large portions of technical reports, thereby reducing a team's main contribution to simply interpreting someone else's research, is clearly not the solution.

Three uses of existing research are common to most technical reports:

- To chronicle the events leading to the approach taken in the current paper and to help the reader understand the context or domain of the problem. This action is typically accomplished in an Introduction or Background section.
- To identify and justify technical parameters needed for the new approach. For the FAA problem, some of these parameters could have been the average airspeed of a Boeing 747 or the typical work hours of an air traffic controller.
- To compare the graphical, symbolic, or numerical results generated by the new modeling approach with those previously identified, so as to examine the benefits or drawbacks of the new approach.

Credible existing research used in these ways does not replace or dilute the current effort but directly supports and strengthens it.

Given the time pressure of the MCM, one has to be very cautious not to get trapped into adopting a complicated modeling component from existing research without being able to explain clearly its development, its use and limitations, and its impact on the current model. This remains the classic red herring of the MCM, luring teams into committing to an approach only to discover late in the process that they are ill-equipped to handle it properly. Ultimately, the evidence of this error appears in the MCM entry in such forms as miraculously appearing formulae, unexplained graphics, and tables of data still waiting to be analyzed. Just as in a court of law, the MCM judges consistently find the results of models built on such tenuous foundations difficult to believe.

About the Author

Pat Driscoll is an Academy Professor of Operations Research in the Dept. of Mathematical Sciences at USMA. He received his M.S. in both Operations Research and Engineering Economic Systems from Stanford University, and a Ph.D. in Industrial & Systems Engineering from Virginia Tech. He is currently the program director for math electives at USMA. His research focuses on mathematical programming and optimization. Pat is the INFORMS Head Judge for MCM/ICM contests.

Practitioner's Commentary: The Outstanding Air Traffic Control Papers

Jack Clemons
Senior Vice President
Lockheed Martin Air Traffic Management
Rockville, MD
`Jack.Clemons@lmco.com`

Introduction

The problem of introducing computer-assisted aids to the air traffic controller community to improve safety and reduce workload is both relevant and immediate. The Federal Aviation Agency (FAA) in the United States, the National Air Transport Services (NATS) agency in the United Kingdom, and numerous other civil aviation authorities around the world currently are engaged in evaluating, developing, testing, and deploying automated support tools even as these MCM papers were being written. All of the papers correctly identified the two factors that determine automation viability: passenger safety and controller workload.

The air traffic controller community is slow to adapt to and embrace new technology. This is not due to some inherent technophobia by this population but rather is driven by professional concern for air traffic safety. Every controller must maintain regular job certification. A serious incident (e.g., a "near miss") that is traced to an error in their performance can cause them to be pulled from their station and recertified. Air traffic control (ATC) is a three-shift-a-day/seven-days-a-week operation. The entire focus of a controller is on the air traffic moving through the airspace as imaged on her display, and the voices of the pilots in her headset. There is little time to have her attention pulled away from those concerns to learn a new, more complicated series of mouse clicks, or to adapt to a new graphic data format or automation aid. The distraction caused is analogous to that of talking on a cellular phone while driving at high speed on the Interstate—attention is diverted from an exclusive focus on safety. And

The UMAP Journal 21 (3) (2000) 305–310.

this is already one of the most stressful professions on earth. The introduction of automation to support and expand the controller's focus must be done with great care.

Criteria: Light from the Real World

Each of the four selected MCM papers addressed the problem assigned in a different way. As a practitioner, I chose not to evaluate each on the elegance of the mathematical modeling employed; the MCM judges can concentrate on that. Rather, I examined each from its relevance to the problems facing real air traffic controllers. This is not to criticize any of the teams for not having the insight and experience of practitioners; of course, that is not realistic. Instead, my intent is to shed some light from the real world on each approach that might be useful to both the team members and to other readers. I assessed each paper on three qualities:

- Thoughtfulness: How well did they think through the problem statement before addressing it?
- Realism: How close does the proposed solution come to addressing a real-world problem?
- Usefulness: Is the proposed solution itself applicable to address the world of the air traffic controller?

On this basis, there was substantial variation in the approaches taken by the four teams.

The Best of the Four: University of Colorado Entry

The entry from the University of Colorado team evaluated best by these measures. It was clear that this team spent appreciable time understanding the FAA. The selection of the Denver International Airport as the system model and obtaining the relevant airport, airspace, and ATC parameters from this airport were done extremely well. The team accurately represented the details of the air traffic and correctly asserted that conclusions drawn should translate into other regions. Second, the team used the Federal Aviation Regulations (FARs) as the guidance for resolving technical assumptions. Once again, this shows a real understanding of FAA procedures. FARs are indeed the governing standard for assessing safety.

The use of a "corrected random walk" to validate the FAA's minimum aircraft separation standard was clever. It ignores standard approaches like the computation of minimum maneuver time required for two aircraft approaching head on. However, the model used by this team provides an insight for parallel flight that is novel and confirms the FAA separation standard as well.

The approach taken to quantify stress as a parameter was good. Once again, they researched the available literature to develop a baseline for ATC stress and built the model from there. The model used is both simple and probably applicable. Though arbitrary, the measure of complexity should demonstrate a logical relationship between traffic patterns and workload stress that can provide insight into underlying causes. Of course, the motivation was to allow the team to transform the automation problem into a problem of queueing, thus greatly simplifying the model and still providing insights. The simple queueing model that the team proposed overlooked a significant factor in runway access, which is gate availability (i.e., even though all runways are clear, traffic may still have to hold because there are no available passenger gates). However, their model can easily accommodate this factor.

The results of this model are intuitive and appear to be correct. Further, for the problem of airport terminal approach and landing at least, they correctly identify which elements of ATC need remain under the control of humans and which have potential for automation.

The Other Outstanding Papers

Each of the other three papers, while providing useful insights, misses this mark by several degrees.

The Duke University Entry

The team from Duke University scored well on my (admittedly subjective) scale of "thoughtfulness." They paid attention to the problem description and showed understanding in the model setup and the trades they performed. However, the paper seems to rely heavily on an assumption that pilots control the aircraft and controllers principally intervene to avoid collisions. This ignores the active control of all aircraft at all times by the controller team as each craft navigates through the airspace.

Commercial airline pilots (as contrasted with general aviation or "private" pilots) do not dynamically choose their routes, their airspeed, or their altitudes. All this is under the direct control of the air traffic controller. Each aircraft files a flight plan, which must be coordinated and cleared with the FAA prior to take off. This flight plan sets the cross-country route that that aircraft will follow, and must explicitly follow regulated "highways" leading from one ATC center to another to ensure continuous ATC coverage. The aircraft's flight is under 100% control of the air traffic controllers from the moment of gate departure through gate arrival.

While the Duke University team's approach was thorough in defining the various "automation" programs that might be employed for collision detection and avoidance, it did not address or encompass the information available to every controller about the planned trajectory of the aircraft obtained from its flight plan. Collision detection schemes currently in use, and more advanced techniques now being evaluated by the FAA, combine information from the radar "track" history for the aircraft with the projected trajectory derived from the flight plan. In addition, the team did not adequately explore the cascade effects of ATC actions. In other words, once a controller decides to have an aircraft change altitude, speed, or direction to avoid a potential conflict, that action must be evaluated to see if new conflicts with other craft have been created. Collision avoidance schemes must have a "what-if?" capability embedded to allow controllers to evaluate a number of potential maneuvers for conflict avoidance before making giving final direction to the pilot. Such techniques are currently under evaluation by NATS in the United Kingdom for implementation in their newest En Route Control Center. On balance, however, I thought this paper had a good approach to realism and usefulness.

The U.S. Military Academy Entry

The approach of the team from the U.S. Military Academy addressed the availability of flight data but appeared to overlook radar. While the mathematical modeling employed seemed sophisticated, the effort was directed toward determining the likely deviation of an aircraft from its flight plan by statistical methods alone. This is both complex and unnecessary, since every aircraft under ATC control can be uniquely tracked by radar. Not only does a controller have a positive radar image (or "track") on his screen, but that track is identified graphically with the aircraft identifier available from the aircraft's transponder. Coupled with the flight plan data available for the aircraft, the ATC has explicit knowledge of position, history, and intent. The statistical model employed by this team is largely irrelevant in that case.

Also, the team makes an assumption that curvature of the earth need not be accounted for, which is not true for high-speed aircraft unless projecting conflicts over short time periods. A conflict model currently being evaluated by the FAA explicitly includes earth curvature for that reason. A similar set of comments applies to the team's complexity model; it is an interesting mathematical exercise but not useful. A controller simply viewing his radar display can readily determine complexity of the airspace. An approach to quantify this complexity for evaluation of automation should be direct, as the simplified approach by the University of Colorado team demonstrates by example.

Virginia Governor's School Entry

Finally, the team from the Virginia Governor's School takes an approach that I believe is the least "real world":

- The approach to validating minimum safe separation was unduly complex.

The team's intent was to attempt to model the aerodynamics of a commercial airliner and, from that, to determine how close another craft could approach before being adversely affected by the air flow. This work has already been done by aircraft manufacturers themselves (and by countless generations of Aeronautical Engineering students, myself among them) and could be readily accessed rather than derived from first principals.

- Once the team set out on this endeavor, they were then forced to make a number of oversimplifying assumptions (e.g., modeling the aircraft as a cylindrical ring, defining the fluid dynamics using vortices and Bernoulli's equation). These assumptions are quite wrong for the model and therefore cause the derived conclusions to be suspect.
- Similar sets of simplifying and unrealistic assumptions were used to model potential conflict. For example, the model takes no consideration of the actual sectors of air space and the active control from ATC under which every aircraft operates.

I believe this team's paper would need substantial rework before it could be credibly presented to the FAA Administrator.

Summary

Each team took a very different approach from the others in addressing the stated problem, with significant variation in practicality from the viewpoint of a practitioner. Of the four papers, that from the University of Colorado would be the one I would recommend taking forward as is to Ms. Garvey and her staff for review and further evaluation.

Reference

Galotti, Jr., Vincent P. 1998. *The Future Air Traffic Navigation System (FANS): Communications Navigation Surveillance Air Traffic Management*. Ashgate Publishing Company, Ltd.

Machol, Robert E. 1995. Thirty years of modeling midair collisions. *Interfaces* 25 (5) (September-October 1995): 51–72.

About the Author

Jack Clemons is the Senior Vice President of Strategic Programs at Lockheed Martin's Air Traffic Company, located in Rockville, Maryland. Jack joined the Lockheed Martin Corporation in April 1996.

Jack began his career at General Electric Corporation's Reentry Systems Division in Valley Forge, Pennsylvania, now part of Lockheed Martin Management and Data Systems. He then worked on the NASA Apollo and Skylab programs for TRW Systems Group in Houston, Texas, and on the NASA Space Shuttle program for IBM Federal Systems Group, also located in Houston. Following that, Jack spent eight years in new product market development and market support for the IBM Corporation in White Plains, New York, and he served a one-year chair assignment as instructor at IBM Corporation's New Management School in Armonk, New York.

Jack joined IBM Federal Systems Group's Air Traffic Control Company in 1992 as Functional Manager of Software Development. Following the acquisition of Federal Systems by Loral, Jack performed the roles of Director of EnRoute Programs and Vice President of Air Traffic Control Engineering. Following the acquisition of Loral by Lockheed Martin, Jack became Senior Vice President of Engineering, Technology and Operations before moving into his current position.

Jack graduated from the University of Florida with Bachelor of Science and Master of Science degrees in Aerospace Engineering.

A Channel Assignment Model: The Span Without a Face

Jeffrey Mintz
Aaron Newcomer
James C. Price
California Polytechnic State University
San Luis Obispo, CA

Advisor: Thomas O'Neil

Introduction

We were asked to design efficient assignment of radio channels to a symmetric network of transmitter locations over a large planar area, so as to avoid interference. Efficiency is based on the span, the minimum of the largest channel assigned.

We derive properties implied by the first set of constraints and by the geometry of the given figure, which we use to construct what we call "span theory." We prove upper and lower bounds for the span of the given figure. With the aid of a computer program, we narrow the bounds and prove that the span is 9. This is also the span of a network generated by extending the figure arbitrarily far in all directions.

We then consider slightly altered constraints, that the channels of neighboring transmitters cannot differ by less than k. We determine two distinct strategies for channel assignments and two associated formulas for the span; the span is $\min\{3k+3, 2k+7\}$ for both the figure and the generated plane.

Allowing a transmitter to be positioned irregularly in the hexagons changes the span by at most 1. Allowing all transmitters to be positioned irregularly—a worst-case scenario—gives a span of 18.

Assumptions and Justifications

- *Every hexagon in the field has a single transmitter at its center*. This can be assumed for Requirements A, B, and C from statements in the problem.

The UMAP Journal 21 (3) (2000) 311–326.

- *Every transmitter is an ideal transmitter*, that is, it transmits radio signals equally well in all directions. No information is given suggesting any of the transmitters are less than ideal. According to our research, an ideal radio station would perform in this manner [Rorabaugh 1990, 134].
- *Every transmitter in the grid is assigned a single channel* (the problem so states).
- *Every positive integer works equally well as a transmission channel*.

Terms and Definitions

- **Neighbors:** Two polygons with a common side.
- **Network:** A finite or infinite group of connected, non-overlapping, maximally packed, regular polygons.
- **Span:** The minimum, over all assignments satisfying the constraints, of the largest channel used at any location.
- **Symmetric network:** A network that is symmetric about some axis.
- **Tessellate:** To repeat a geometric pattern over an infinite or finite plane.
- **Tessellation:** An arrangement of polygons that will fit together without overlapping or creating any gaps and cover an infinite plane.
- **Valid network:** A network that satisfies all the constraints of the given requirement.

Analysis

Hexagonal Geometry

For convenience, we rotate the figure given in the problem to obtain rows instead of columns, as shown in **Figure 1**.

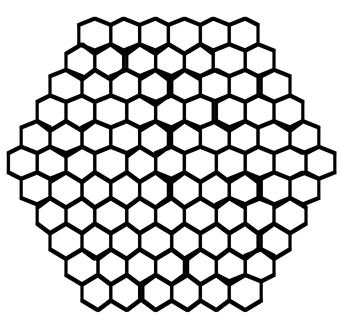


Figure 1. Rotated horizontal layout of hexagons.

We begin by analyzing the geometry of a regular hexagon with side length s. Drawing the diagonals as shown in **Figure 2** yields six regular triangles. The distance from the center to any corner is s, and the perpendicular distance from the center to any side is $s\sqrt{3}/2$.

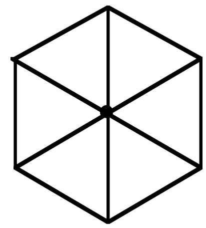


Figure 2. Detail of a regular hexagon.

We examine a small symmetric network to determine the effects of the spectral spreading constraint. **Figure 3** illustrates a circle of radius $2s$ drawn from the center hexagon. The centers of all six adjacent hexagons are within this circle, so no transmitter may neighbor a transmitter of an adjacent channel; but the circle does not spread beyond these six hexagons, so any transmitter beyond the six hexagons that surround the center may be assigned to an adjacent channel without interference.

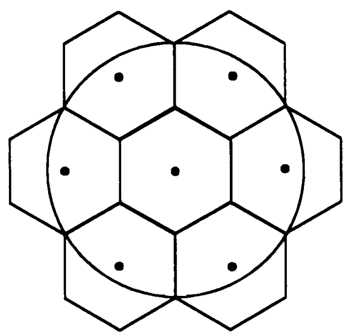


Figure 3. A symmetric network with seven hexagons and radius $2s$ marked from the center of the center hexagon.

For example, consider **Figure 4**, with two networks in which channels are assigned to each hexagon. The network on the left violates the constraint above, because channel 1 and channel 2 are neighbors; the network on the right is a valid network.

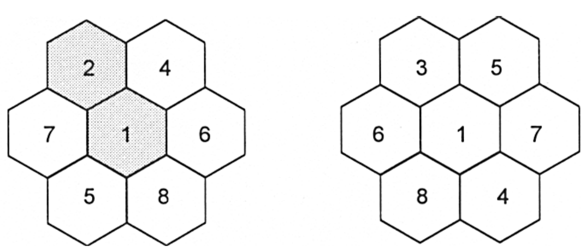


Figure 4. Two symmetric networks with seven hexagons and assigned channels.

We summarize the effect of spectral spreading with the following rule:

Adjacent Channel Principle: *No neighboring hexagons may be assigned adjacent channels.*

We now consider the requirement that no two transmitters with the same channel may be within $4s$ of each other.

We define two hexagons to be *n hexagons away from each other* if and only if one can construct a path of n straight line segments, and no fewer, with each segment having length $s\sqrt{3}$ and having both endpoints at centers of hexagons. For example, in **Figure 5**, hexagon A is 3 hexagons away from hexagon B.

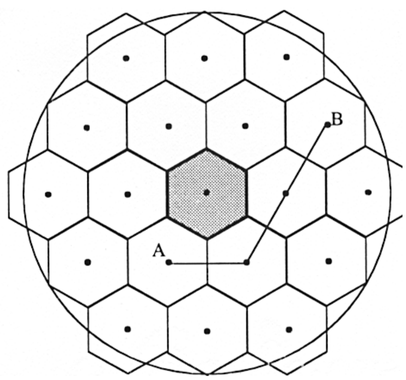


Figure 5. Hexagon A is 3 hexagons away from hexagon B.

From **Figure 5**, we observe that no two transmitters of the same channel can be 2 hexagons away from each other, but they can be 3 hexagons away from each other:

Same Channel Principle: *No two transmitters of the same channel may be less than three hexagons away from each other.*

Embedded Subgraph Method

We divide the hexagons into sets as indicated in **Figure 6**. The dotted lines indicate the embedded subgraph within the network. The vertices of the subgraph are indicated in the hexagons that are included in the subgraph, and those hexagons are marked as well. The vertices are intentionally drawn off-center so that the subgraph's edges do not coincide with any hexagon sides. The shaded hexagons alternate blue and green.[1]

The model works as follows. Any blue hexagon is at least 3 hexagons away from any other blue hexagon; so by the **Same Channel Principle**, all blue hexagons may be assigned the same channel. Similarly, all green hexagons may

[1] EDITOR'S NOTE: Like **Figure 6**, a number of the authors' original figures are in color, and readers of the electronic version of this article can see them in color. Expense prohibits reproducing the figures as color plates in the printed *Journal*.

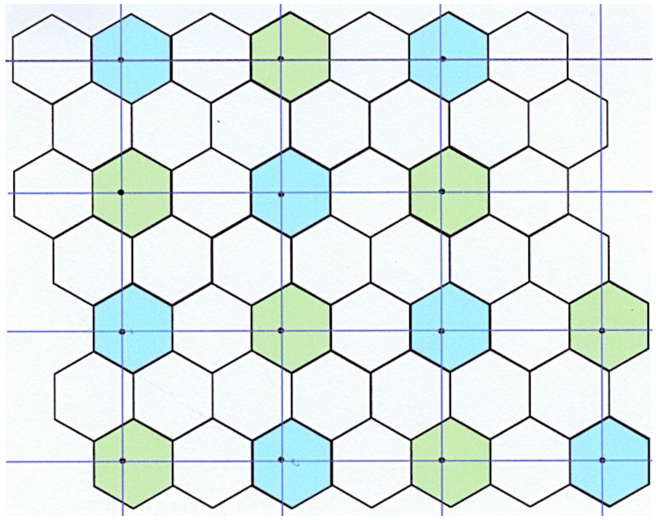


Figure 6. Embedded Subgraph Method setup.

be assigned the same channel. No blue hexagon neighbors any green hexagon; by the **Adjacent Channel Principle**, we may assign blue and green hexagons adjacent channels.

There are four such embedded subgraphs in the network. The hexagons on the same rows as the blue and green hexagons, but between them, create another embedded subgraph, which can also be assigned two adjacent channels. Next, we can place two more subgraphs to connect the rows that have not been assigned channels yet. The result is four embedded subgraphs that together contain all the hexagons in the network. **Figure 7**, in which the first and third rows alternate colors and the second and fourth rows alternate different colors, shows how every hexagon is assigned to exactly one subgraph. Hexagons of the same pattern belong to the same subgraph.

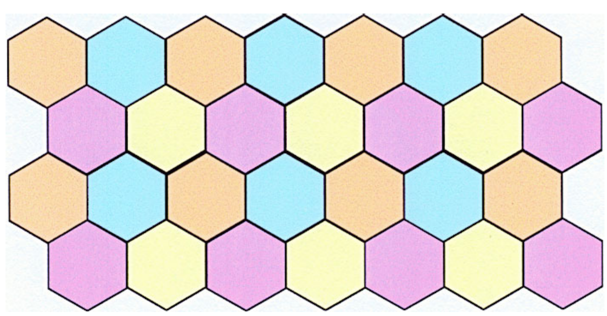


Figure 7. Embedded subgraphs indicated by shading.

The final step is to assign channels to all four subgraphs with two channels per subgraph, as described above. For the first subgraph, we use channels 1 and 2. At this point every hexagon in the other three subgraphs neighbors a channel 2 hexagon; thus no hexagon may be assigned channel 3, by the **Adjacent Channel Principle**. So we skip 3 and assign 4 and 5 to the hexagons of a second subgraph (it does not matter which subgraph). As before, each remaining hexagon neighbors a 5; so we skip channel 6 and assign 7 and 8 to a third subgraph. Finally, we skip channel 9 and assign channels 10 and 11 to the remaining subgraph. This pattern yields three channel values not being assigned. The resulting pattern is shown in **Figure 8**.

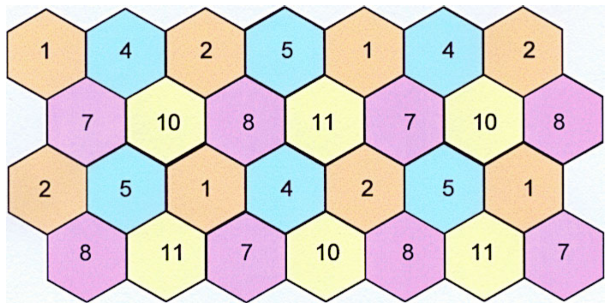


Figure 8. Channels assigned using the Embedded Subgraphs model.

For the sake of examining interference, the distance to transmitters of the same channel is at least $s\sqrt{21}$. The distance to adjacent channels is at least $3s$. Also, every channel has only one adjacent channel that appears in the network.

This pattern can be tessellated across any number of hexagons. Thus, we know that the span is at most 11 for **Figure 1** as well as a grid that spreads arbitrarily far in all directions.

Diagonals Method

We attempt to fill the network using diagonals. We want exactly three channels repeated along any diagonal. To avoid same-channel interference, each channel along the diagonal should be exactly three hexagons away from the closest hexagon with the same channel. This diagonal should then be repeated three diagonals away on either side. We call such repeated diagonals *same diagonals*. The result is shown in **Figure 9**. The channels x, y, and z are chosen so that no two of them are consecutive channels and spectral spreading is avoided. For example, x, y, and z may be assigned the values 1, 3, and 5.

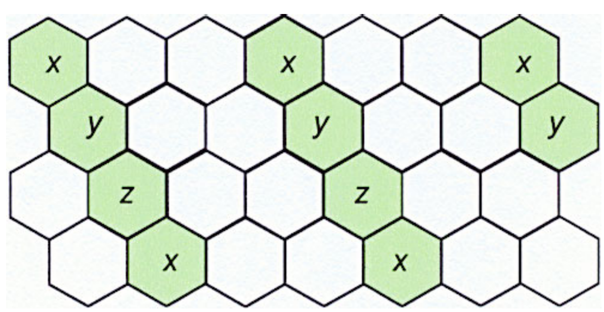


Figure 9. Diagonals Method setup.

After assigning channels to the first diagonal, we attempt to assign values to its neighbor diagonals without violating any constraints. With trial and error, we obtain the assignments shown in **Figure 10**.

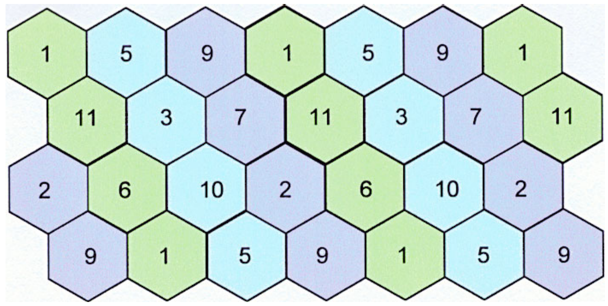


Figure 10. Channel assignments made using the Diagonals Method.

This tessellation also has a maximum channel of 11, and like the Embedded Subgraphs Method, it can be expanded infinitely in all directions. Also, the distance to the closest adjacent channel is $3s$. Unlike the Embedded Subgraphs Method, this pattern contains nine unique channels, leaving two unused (channels 4 and 8). Furthermore, the distance to the closest transmitter with the same channel is $3s\sqrt{3}$. Another difference from the Embedded Subgraphs Method is that three of the channels have two adjacent channels appearing in the network, while six of them have only one.

We explain why the assignment in **Figure 10** is valid. Consider every hexagon as belonging to one of three subgraphs, whose edges create a set of triangles. With the 1, 2, and 3 placed as they are, they violate no rules; but placing a 4 in any hexagon would violate the **Adjacent Channel Principle**. So we skip 4 and assign channels 5, 6, and 7 to a triangular subgraph with the same properties as the first subgraph. Finally, we skip 8 and assign 9, 10, and 11 to the remaining subgraph.

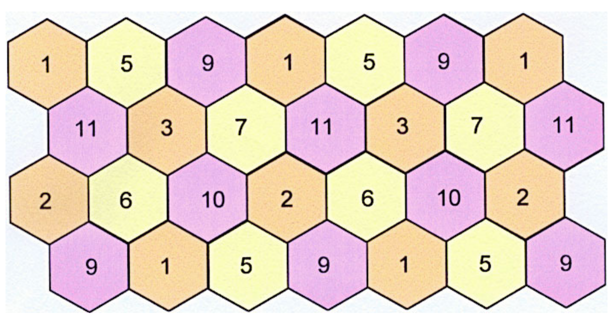


Figure 11. Embedded subgraph model setup.

Computer Program

We confirmed our results (maximum channel of 11) by writing a computer program to test every possible combination and determine the span. (Before writing the program, we knew that it would require 91 levels of recursion, but the speed and memory capacity of modern PCs makes this a feasible option.)

The program begins with an empty network, that is, a network with no channels assigned to any transmitter. It starts with the farthest left hexagon in the top row of the network and assigns it channel 1. It moves across the row and then similarly processes the following rows. Each time a new hexagon is encountered, the program attempts to assign a channel to the hexagon, starting with 1 and ending at some user-specified upper bound. After assigning a channel, it checks that the assignment does not violate the **Adjacent** and **Same Channel Principles** listed above. If there is a violation, it tries the next highest channel. If the channel assigned exceeds the upper bound, the program moves back to the previous hexagon and performs a reassignment before continuing. If the program successfully assigns a value to the last hexagon in the last row, it has successfully filled the network and displays the entire network for the user. If a network is never displayed, the upper bound was too low.

The program output various working configurations for an upper bound of 9 but no layouts for 8. The pattern that generates a span of 9 is shown in **Figure 12**. This pattern appears to be a variation of the Diagonals Method: the hexagons of any right diagonal consist of exactly three channels, which are repeated ad infinitum. However, as highlighted in the figure, for any entry in a diagonal, the corresponding entries of the closest same diagonals are on different rows. This does not change the fact that the **Same Channel Principle** is satisfied; any pair of same-channel transmitters are still at least three hexagons away from each other.

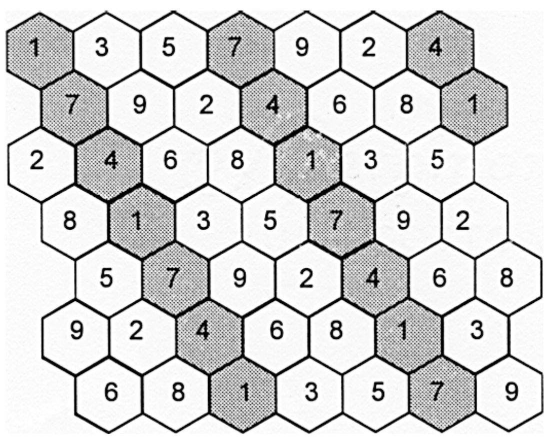


Figure 12. Computer-generated span of 9.

To verify that this is a valid pattern, we must also show that the **Adjacent Channel Principle** is not violated. The channel of any hexagon differs from its left and right neighbors by at least two. Every row is identical to the one above it but shifted three diagonals to the left. Due to the shift, neighbors above and below differ by at least three. Thus, the **Adjacent Channel Principle** is satisfied.

We develop "span theory" to prove that the span of **Figure 1**, as well as the span of an arbitrarily large grid, is 9.

Span Theory

Let S be the set of all networks constructed with regular hexagons of side length s, and let H_n be the symmetric network of regular hexagons with n regular hexagons on any side of it (**Figure 13**). Note that **Figure 1** is H_6; we denote by H_∞ the symmetric network that spreads arbitrarily far in all directions.

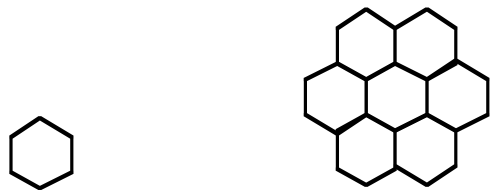


Figure 13. Symmetric networks H_1 and H_2.

Let $A \in S$. We define Span(A) as the minimum of the largest channel used at any location in A, over all channel assignments to the hexagons in A satisfying both the **Adjacent** and **Same Channel Principles**. From our previous results, we have:

Maximum Span Property: *For all $A \in S$, Span$(A) \leq 9$.*

Our goal is to show that Span(A) is also bounded below by 9 for $A = H_6$ and $A = H_\infty$.

Let $A, B \in S$. We define $A \subset B$ if A can be traced out entirely inside B, that is, B can be truncated by removing hexagons so that it is congruent to A.

Proposition 1. *If $A \subset B$, then Span$(A) \leq$ Span(B).*

[EDITOR'S NOTE: We omit the proof.]

Lemma 1. *If the channel of the center transmitter in an H_2 network is not 1 or 8, then there is a channel greater than or equal to 9 in that network.*

Proof: Suppose not. Let x be the channel of the center hexagon, $1 < x < 8$. Then there are at most 8 choices for the seven hexagons in the H_2 network. Let T be set of possible channels for the six neighbors of the center hexagon. By the **Same Channel Principle**, $x \notin T$. Since $x > 1$, we have $x - 1 \geq 1$, so by the **Adjacent Channel Principle**, $x - 1 \notin T$. Similarly, since $x < 8$, we have $x + 1 \leq 8$, and by the **Adjacent Channel Principle**, $x + 1 \notin T$. So $T \subset \{1, 2, 3, 4, 5, 6, 7, 8\} - \{x - 1, x, x + 1\}$, so $|T| \leq 5$. Since we must assign to the six hexagons at most five different channels, at least one channel is assigned to two hexagons. Doing so violates the **Same Channel Principle**, so we arrive at a contradiction. $\square$

Theorem 1. Span$(H_6) = 9$.

Proof: Let A be as shown in **Figure 14**; $A \in S$.

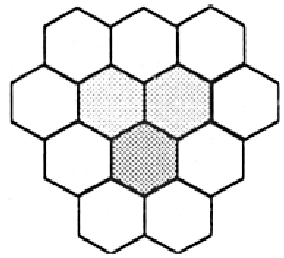


Figure 14. Network A of Theorem 1.

Note that A has three hexagons (shaded in the figure) that are the centers of three corresponding H_2 networks. By the **Same Channel Principle**, since these three hexagons are neighbors, they must be assigned three unique channels. Thus, at most one can be assigned channel 1, at most one can be assigned channel 8, and therefore at least one of the shaded hexagons must be assigned a number other than 1 or 8. Since the shaded hexagon with a channel other than 1 or 8 is the center of an H_2 network, that network has a channel greater than or equal to 9 in it, by **Lemma 1**. Since A contains this H_2 network, there must be a hexagon in A with a channel greater than or equal to 9 in it. Thus, $\text{Span}(A) \geq 9$. By the **Maximum Span Property**, $\text{Span}(H_6) \leq 9$. By **Proposition 1**, since $A \subset H_6$ we have $9 \leq \text{Span}(A) \leq \text{Span}(H_6) \leq 9$. Therefore, $\text{Span}(H_6) = 9$. □

Corollary 1. $\text{Span}(H_\infty) = 9$.

Requirement C

We move to the case in which transmitters within distance $2s$ differ by at least some given integer k. The result of this new constraint is a modified principle for adjacent channels:

The k-Adjacent Channel Principle: *No neighboring hexagons may have channels that differ by less than k.*

Requirement C states that the distance to same-channel transmitters remains at least $4s$, so the **Same Channel Principle** applies as before.

We must modify our definition of $\text{Span}(A)$: Let $A \in S$ and $k > 1$. We redefine $\text{Span}(A, k)$ as the minimum of the largest channel used at any location in A, over all channel assignments to the hexagons in A that satisfy the k-**Adjacent** and **Same Channel Principles**. Thus, $\text{Span}(A) = \text{Span}(A, 2)$.

Note that the analogue of **Proposition 1** holds, that is, if $A \subset B$, then $\text{Span}(A, k) \leq \text{Span}(B, k)$.

We want to find $\text{Span}(H_6, k)$ for general k. Running our computer program to find networks satisfying the k-**Adjacent Channel Principle** with various values of k, we find

$$\begin{array}{lll} \text{Span}(H_6, 2) = \ \ 9 & \text{Span}(H_6, 7) = 21 & \text{Span}(H_6, 3) = 12 \\ \text{Span}(H_6, 8) = 23 & \text{Span}(H_6, 4) = 15 & \text{Span}(H_6, 9) = 25 \\ \text{Span}(H_6, 5) = 17 & \text{Span}(H_6, 10) = 27 & \text{Span}(H_6, 6) = 19 \end{array}$$

So we hypothesized that

$$\text{Span}(H_6, k) = \begin{cases} 3k+3, & 2 \le k \le 4; \\ 2k+7, & k \ge 4. \end{cases}$$

We found two distinct patterns, one with maximum channel $3k+3$ and the other with maximum channel $2k+7$ (**Figure 15**).

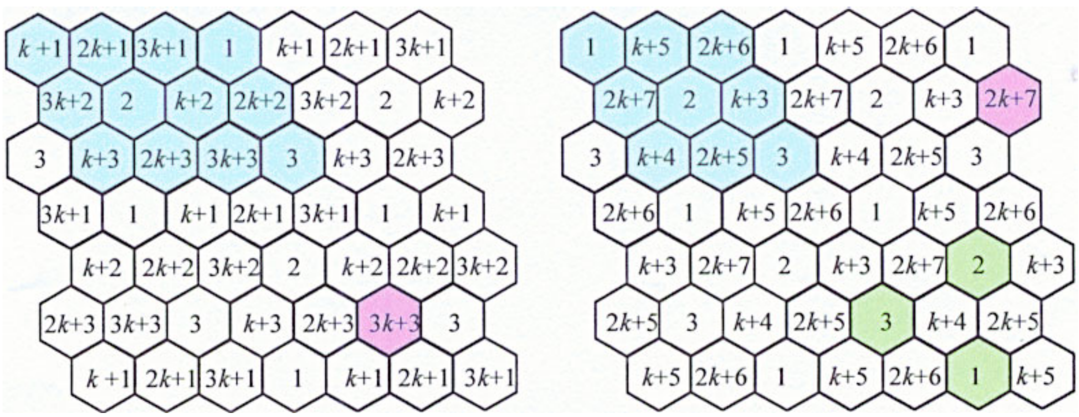


Figure 15. Patterns for channel assignments.

Highlighted with shading in the upper left-hand corner of each pattern is the tessellation that generates the network. Also shaded is the maximum channel used by each tessellation. The $2k+7$ pattern is actually an arrangement that fits the Diagonals Method; a sample triangular subgraph is highlighted in the lower right-hand corner.

The $2k+7$ pattern uses channels $1, 2, 3, k+3, k+4, k+5, 2k+5, 2k+6$, and $2k+7$. It is valid because it follows the same rules as the Diagonals Method:

- Specifically, any given channel is three hexagons away from a hexagon with the same channel on the same row. Also, the nearest row with the same channel is three rows away. Thus the **Same Channel Principle** is satisfied.

- We note that any channel cannot be placed next to any other channel in its column for $k > 2$. However, it can be placed next to every entry in the other columns, since they will differ by at least k. Since every entry has six unique neighbors, placing the six channels from the other two columns around the entry results in a valid arrangement. Examining the network confirms this deduction: Every 1, 2, or 3 is surrounded by the six channels listed in the other two columns; the same also holds for these six channels. Thus, the **k-Adjacent Channel Principle** is satisfied.

The verification of the $3k+3$ pattern is similar.

Comparing the two patterns, the minimum distance to a transmitter of the same channel is $3s\sqrt{3}$ in both patterns, and the minimum distance to the nearest adjacent channel is $3s$ in both. From these results, we can state a modified maximum span property:

Second Maximum Span Property. *If $A \in S$, then*

$$Span(A,k) \leq \min\{3k+3, 2k+7\}.$$

Theorem 2. *If $A \in S$, then $2k+4 \leq Span(A,k) \leq \min\{3k+3, 2k+7\}$.*

Lemma 2. *If $k > 6$, then $Span(H_4,k) = 2k+7$.*

Theorem 3. *If $A \in S$, $H_4 \subset A$, and $k > 6$, then $Span(A,k) = 2k+7$.*

[EDITOR'S NOTE: We omit the proofs of these results.]

There are a few cases that our mathematical results do not cover; but since our program verifies all results that we obtain mathematically, we are confident that it can find Span(A,k) for any $A \in S$, $k > 1$.

Requirement D

We begin by considering the case of irregular transmitter placements and analyze two cases.

All Transmitters Except One Are in Hexagon Centers

The exception may be anywhere in its hexagon. How far from the center of a hexagon can a transmitter be and still be in the hexagon? Just s. We consider the constraints of Requirement A.

Adjacent-channel transmitters must be $2s$, and same-channel transmitters $4s$, away from each other. If we give one transmitter freedom to move up to s away from its center, then to avoid interference the distance between the center of the irregularly placed transmitter and the other transmitters must not be less than $2s + s = 3s$ for adjacent-channel transmitters, and $4s + s = 5s$ for same-channel transmitters.

The change to $3s$ has no effect, since there are no transmitters between $2s$ and $3s$ away from the center of a hexagon. Thus, the **Adjacent Channel Principle** does not change. However, there are 12 hexagons whose centers are $s\sqrt{21}$ away, which means they could have been same-channel transmitters before (by the **Same Channel Principle**) but now we cannot be sure. Thus, the **Same Channel Principle** no longer holds, as we have transmitters three hexagons away that cannot have the same channel. These 12 hexagons are the shaded ones in a ring in **Figure 16**. The center hexagon is the one that can be irregularly placed.

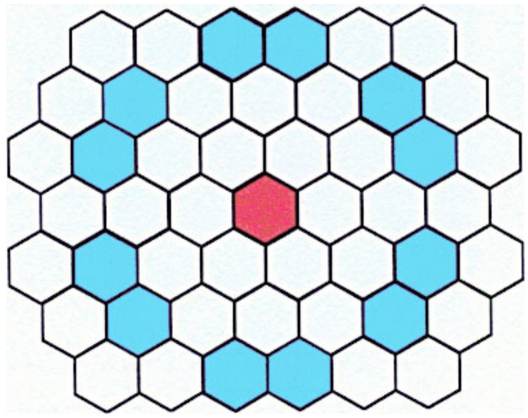


Figure 16. The channel of any shaded hexagon cannot be the same as that of any white hexagon.

Allowing one transmitter to be irregularly placed has minimal effect on the span, changing it by at most 1, a fact that we prove. Let $A \in S$, $k > 1$, and $n \geq 0$. We define Span$_n(A, k)$ as the minimum of the largest channel used at any location in A, over all channel assignments to the hexagons in A that satisfy the k**-Adjacent** and **Same Channel Principles**, with the additional allowance that up to n transmitters in A may appear anywhere within their respective hexagons. Note that Span$_0(A, k) =$ Span(A, k).

Theorem 4. *Span*$_1(A, k) \leq$ *Span*$(A, k) + 1$.

Proof: Let $x =$ Span(A, k). We can assign channels to all hexagons in A so as to yield a span of x, with the irregularly placed hexagon assigned channel x. By the **Same Channel Principle**, all transmitters with channel x are at least three hexagons away from the irregularly placed transmitter. By the k**-Adjacent Principle**, the irregularly placed hexagon has no neighbors with channels greater than $x - k$. Change the assignment of the irregularly placed transmitter from x to $x + 1$. Since there are no other transmitters with channel $x + 1$, the **Same Channel Principle** is satisfied. Furthermore, since there are no hexagons neighboring the irregularly placed transmit with channels greater than $x - k$, the k**-Adjacent Channel Principle** is satisfied. Thus, we have constructed a valid pattern for A with highest channel $x+1$ and a single irregularly placed transmitter. So Span$_1(A, k) \leq x + 1 =$ Span$(A, k) + 1$. □

Any Transmitter Can Occupy Any Position within Its Hexagon

The farthest that a transmitter can be from its center is s. Suppose that two transmitters need to be at least $2s$ apart; to guarantee that they are $2s$ units apart, we require that their hexagons' centers be $2s + 2s = 4s$ apart. To guarantee that transmitters are $4s$ apart, we can place their hexagons' centers at least $4s + 2s = 6s$ apart.

Placing this scenario in our computer program, we find a pattern for the case of the large network (**Figure 17**).

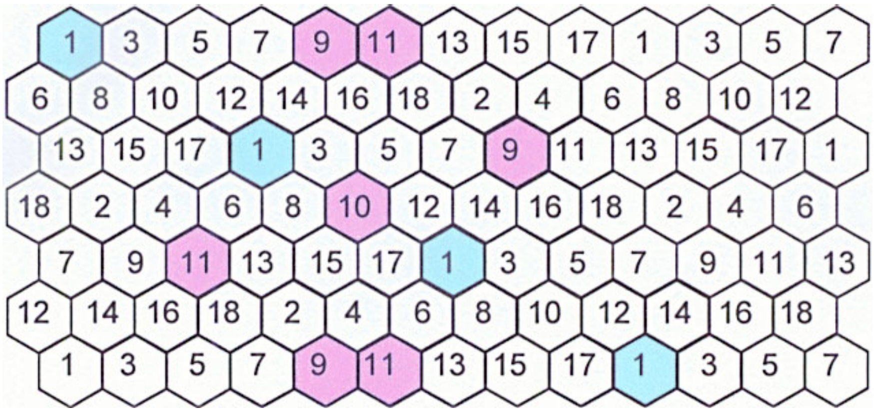


Figure 17. A valid network in which any transmitter can be located anywhere in its hexagon.

This network has a maximum channel of 18. Every channel is the minimum distance from a same-channel transmitter on two sides, as illustrated by the shaded hexagons assigned channel 1. Every transmitter is the minimum distance from each adjacent channel on three sides; that is, there is hexagon closer that could be assigned an adjacent channel and still satisfy both principles. If a channel is not 1 or 18, it has two adjacent channels, so it is the minimum distance from six transmitters with adjacent channels; this is illustrated by the shaded hexagon assigned channel 10, which is the minimum distance from the six shaded hexagons assigned channels 9 and 11.

Since all 18 channels are used, each transmitter is minimally close to at least five other transmitters, and since our program cannot find a pattern with maximum channel 17, we conjecture that Span$(H_6, 2) =$ Span$(H_\infty, 2) = 18$.

Several Levels of Interference and Other Factors

We were asked to consider generalizations of the problem.

- A network with several levels of interference could imply a third level of interference, in addition to same-channel interference and spectral-spreading interference. It could also imply varying levels of spectral spreading, such as a rule that channels differing by 1 must be at least $3s$ apart, channels differing by 2 must be $2s$ apart, and same-channel interference remains unchanged.

- Interference levels may vary in different parts of the grid. For example, the top half of the network may satisfy the **2-Adjacent Channel Principle** while the bottom half requires the **3-Adjacent Channel Principle**. Our program could be modified to calculate the span of such a network.

- Transmitters could have non-repeated channels, or there may be certain channels that no transmitters may use.

- Perhaps a small amount of spectral spreading is acceptable, that is, a network might set n as the limit to the number of transmitters allowed within distance

$d \cdot s$ with channels differing by less than k, for some nonnegative integers d, k, and n. In Requirement A, spectral spreading would be described by $n = 0$, $d = 2$, $k = 2$. The **Same Channel Principle** would be defined by $n = 0$, $d = 4$, $k = 1$.

As the problem states, one basic approach is to partition the region into regular hexagons. Squares and triangles could also be used; these are the only regular polygons besides hexagons that tessellate a plane [Firby and Gardiner 1982, 151]. Our program could easily be modified to handle such networks.

Press Release

Radio Channel Assignments Problem Solved:

Robust Computer Program, Mathematical Theory Pave Path to Solution

SAN LUIS OBISPO, Feb. 30 — A team of three undergraduates students from CalPoly cracked the case of the radio channel assignments. The team had to assign channels to radio transmitters on a hexagonal grid in such a way as to prevent several levels of frequency interference.

The team first determined that no more than 11 channels would be needed; then a computer program suggested that a solution with 9 channels might be possible. To prove that 9 channels—but no fewer—would work, the team developed "span theory," a new mathematical theory of channel assignment.

The team also solved several more general problems accounting for wider channel separation or allowing transmitters to be moved around.

Strengths and Weaknesses

Through a series of models and some mathematical theory, we find and rigorously prove that the span of the given figure in Requirement A, and an arbitrarily large figure in Requirement B, is 9.

Our computer program verifies all of our results and is an invaluable tool for determining patterns. It is very robust in its ability to calculate spans for networks of almost any size, subject to constraints that can easily be modified. Execution is almost instant for networks with fewer than 100 hexagons. It would be very difficult to prove that the code is correct, but we develop a rigorous span theory to prove the values of spans.

We also present some early heuristics, the Embedded Subgraph Method and Diagonals Method, that provide near-span solutions and are easily shown to be valid without span theory. In some scenarios, these methods might be preferred for assigning channels, such as if certain channels are forbidden.

For Requirement C, aided by our computer program and span theory, we find two strategies and set upper and lower bounds on the span of a network. With span theory, we find the exact value of the span both for **Figure 1** and for an arbitrarily large network, as well as for hexagonal networks with side length exceeding 3, provided $k > 6$. Though we cannot rigorously determine the spans for $3 \leq k \leq 6$, the computer program calculates them.

We consider only extreme cases of irregular placement in Requirement D. We prove a formula for the span if only one transmitter is allowed to be irregularly placed. On the basis of our program, we also find a span, 18, for allowing every transmitter to be irregularly placed. Channel assignments are valid no matter where a transmitter is moved, provided that it stays within its hexagon.

References

Conway, John H., and Richard Guy. 1996. *The Book of Numbers*. New York: Springer-Verlag.

Firby, P.A., and C.F. Gardiner. 1982. *Surface Topology*. Chichester, UK: Ellis Horwood.

Johnsonbaugh, Richard. 1993. *Discrete Mathematics*. 3rd ed. New York: Macmillan.

Rorabaugh, C. Britton. 1990. *Communications Formulas and Algorithms*. New York: McGraw-Hill.

Saaty, Thomas and Paul Kainen. 1977. *The Four Color Problem: Assaults and Conquests*. New York: McGraw-Hill.

Smith, Douglas, Maurice Eggen, and Richard St. Andre. 1997. *A Transition to Advanced Mathematics*. 4th ed. Monterey, CA: Brooks/Cole.

"We're Sorry, You're Outside the Coverage Area"

Robert E. Broadhurst
William J. Shanahan
Michael D. Steffen
Lewis and Clark College
Portland, OR

Advisor: Robert W. Owens

Our Approach

We assume that the physical transmission properties do not result in penetration or interference varying from channel to channel. Since the channels occupy a continuous portion of the frequency spectrum, they can be numbered with integers from 1 up to the number n of channels; n represents the bandwidth of the portion of the spectrum. The minimum possible value for the bandwidth that achieves all the requirements for a given transmitter arrangement we call the *span* of that arrangement.

Our problem is to provide a method for arranging channels among the transmitters that achieves as low a bandwidth as possible. We analytically establish a lower limit for the span and we find the best solutions that we can to the given problem.

The bandwidth of a feasible solution is an upper limit on the span. We seek to raise (through further analysis) the lower limit and to lower (through further construction) the upper limit. If our upper limit meets our lower limit, then we have completely determined the span.

Requirements A and B are special instances of the more general problem posed in Requirement C; we treat A, B, and C as one main problem and solve them together. Requirement D asks us to consider generalizations of the problem; we examine how to treat some of the weaknesses in our main problem. Finally, Requirement E asks us to write an article for the local newspaper, which appears at the end.

The UMAP Journal 21 (3) (2000) 327–342.

The Main Problem

At any point in the plane, a receiver can obtain a clear signal from transmitters within some maximum range, without interference from the signals broadcast by neighboring transmitters.

Assumptions

- The region of interest is partitioned into adjacent regular hexagons of the same size.
- The length of the side any hexagon is s.
- Each hexagon represents the area serviced by one transmitter, which is located in the center of its hexagon.
- Each transmitter broadcasts a single channel.
- To minimize interference, any two transmitters occupying the same channel must be at least $4s$ apart.
- To minimize sideband interference, transmitters less than $2s$ apart must use channels that differ by at least k channels. (In Requirements A and B, $k = 2$.)
- The region of interest is a field of indefinite size and shape. A specific region is given for Requirement A and for a special case of Requirement C.

Definitions

- **Cell:** The area serviced by a given tower—in this case, hexagons.
- **Distance:** The minimum number of edges that must be crossed to move from one cell to another. Any hexagon in an infinite field is surrounded by six cells of distance 1, twelve of distance 2, and so on.
- **Separation:** The absolute value of the difference between the integers assigned to two channels.

The Model

Each hexagon is labeled with the integer channel of its transmitter. To satisfy the $2s$ requirement, adjacent hexagons must have a channel separation of at least k. To satisfy the $4s$ requirement, neither adjacent hexagons nor hexagons that are a distance of 2 away can be labeled with the same integer; they must have a separation of at least 1. In **Figure 1**, the inner ring of 6 unmarked hexagons cannot assume the same label as the center hexagon or labels with a separation of less than k. The surrounding ring of 12 barred hexagons cannot be labeled with the same value as the center.

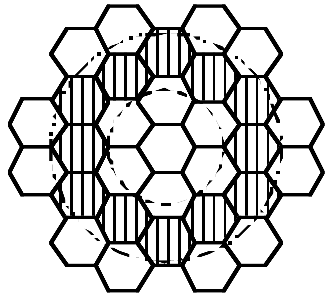


Figure 1. The dashed circle has radius $2s$, the dash-dotted circle has radius $4s$.

We establish a minimal value for the span:

Theorem. *The span of any solution is at least 7.*

Proof: Given any hexagon X, there are 6 additional hexagons whose centers are within a distance of $2s$ from the center of X. Any two of these 7 hexagons have centers that are less than $4s$ apart. Therefore, no 2 of the 7 hexagons can have the same label. □

Theorem. *For $k = 1$, the span is 7.*

Proof: The span is at least 7. A solution with 7 as the largest label used is illustrated in **Figure 2**. A solution for the infinite region of interest can be constructed by tiling the plane with the center group of 7 cells. □

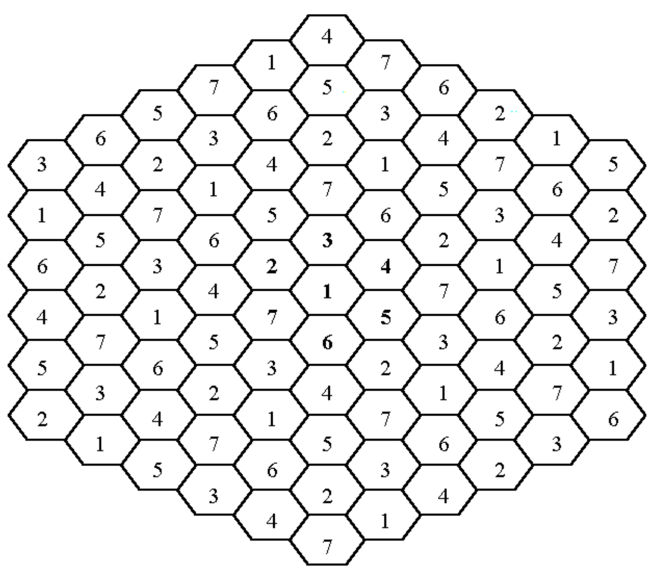


Figure 2. Solution showing that the span is 7.

Theorem. *For any $k \geq 1$, the span is at least $2k + 5$.*

Proof: Select three cells such that each is adjacent to the others and no cell is on the boundary of the region of interest. No two of these three cells can assume the same label. Thus, there must be a maximum, a middle, and a minimum value: A, B, and C respectively. Since two adjacent cells' labels cannot be separated by less than k, we know that $A - B \geq k$ and $B - C \geq k$.

Consider the six hexagons adjacent to B. The minimum separation between B and any adjacent hexagon is k, so the labels of the cells adjacent to B cannot be any of B, $B \pm 1$, $B \pm 2$, ... , $B \pm (k - 1)$. Therefore, there are $2(k - 1) + 1$ channels between A and C that cannot be adjacent to B. There must be at least 6 distinct channels adjacent to B; thus, there must be at least $2(k - 1) + 7$ channels, so the span is at least $2k + 5$. □

Theorem. *For $k = 2$, the span is 9.*

Proof: By the preceding theorem, the span is at least 9. A solution for Requirement A with 9 as the largest label used is illustrated in **Figure 3**. □

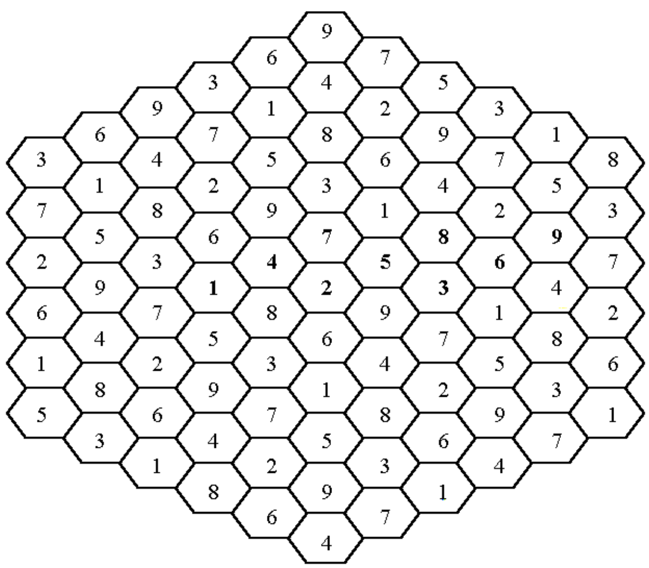


Figure 3. Solution to Requirement A with span 9.

The solution contains a center repeating group of 9 cells that tiles the plane, hence providing a solution to the case in Requirement B. An additional method of generating this arrangement is to add 2 whenever you move from one cell to its neighbor up and to the left, add 3 when you move up and to the right, and add 4 when you move down; if the result is greater than 9, subtract 9. Label each cell with the result and continue in all directions. If you start at a cell labeled x and move once in each direction, you will return to the starting cell. You will have added 2, 3, and 4 to x and subtracted 9 once, giving a net label of x. Therefore, this labeling algorithm is consistent.

At this point, the span has been completely specified for the cases $k = 1, 2$.

Theorem. *For an infinite region of interest and for $k \geq 2$, the span is no more than* $3(k+1)$. (This extends the spanning solution for $k = 2$.)

A solution for $k = 3$ is displayed in **Figure 4**; the group of 12 cells numbered in bold tiles the plane. We do not prove this theorem for general k, because we improve on it shortly.

We summarize in **Table 1** what we know about the span as a function of k.

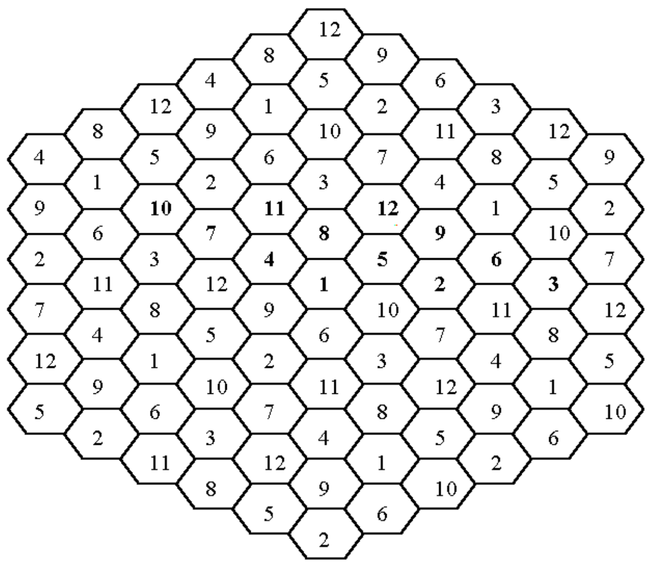


Figure 4. Solution for $k = 3$, with span 12.

Table 1.
Minimum and maximum values for the span by value of k.

k	Minimum	Maximum
1	7	7
2	9	9
3	11	12
4	13	15
5	15	18
k	$2k+5$	$3k+3$ (not proven)

We develop some further theory to help us determine the span for $k = 3$.

Theorem (The Symmetry Argument). *If a solution with span s uses label x, then there is a solution with the same span that uses label $s+1-x$.*

Proof: The cells of the given solution are labeled with values between 1 and s. Relabel each cell y with the label $s+1-y$. Since $1 \leq y \leq s$, then $1 \leq s+1-y \leq s$. The difference $|a-b|$ between the labels a and b of any two cells does not change, because $|(s+1-a)-(s+1-b)| = |a-b|$. Therefore, the new labeling is also a solution, because solutions depend on only the differences between the labels of the cells. We know that at least one cell was originally labeled x; that cell is now labeled $s+1-x$. $\square$

Contrapositive of the Symmetry Argument: *If there are no solutions of span s that include label x, then there are no solutions of span s that include label $s+1-x$.*

Theorem. *For $k = 3$, the span is 12.*

Proof: The span is at least 11. Suppose that there is a solution that uses label 3. Labels 1, 2, 3, 4, 5 cannot be adjacent to 3, so the six adjacent hexagons must be labeled 6, 7, 8, 9, 10, and 11. The only label that could be adjacent to both 3 and 9 would be 6; but we need at least 2 channels that are adjacent to both (**Figure 5**). Hence, 3 cannot be used in a solution with span 11. By the contrapositive of the symmetry argument, no solutions of span 11 include label 9, either.

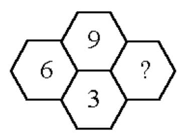


Figure 5. Situation that arises if label 3 is used in a span of 11.

Now suppose that label 4 is used. Then labels 2, 3, 4, 5, 6, and 9 cannot be adjacent to 4, leaving only labels 1, 7, 8, 10, and 11. We need 6 distinct labels adjacent to 4 but only 5 remain. Thus, 4 cannot be used in a solution with span 11. Again by symmetry, no solutions of span 11 include label 8.

An identical argument excludes labels 5 and 7. Only labels 1, 2, 6, 10, and 11 remain; but at least 7 labels are required for a solution. Thus, a solution for $k = 3$ with a span of 11 is not possible. We have already constructed a solution of bandwidth 12 for $k = 3$, so the span is 12. $\square$

We turn to $k \geq 4$.

Theorem. *For $k \geq 4$, the span is no more than $2k + 7$.*

Proof: A solution satisfying all constraints may be constructed as follows. Consider **Figure 6**. Let $A_1, A_2, A_3 = 1, 2, 3$; $B_1, B_2, B_3 = k+3, k+4, k+5$; and $C_1, C_2, C_3 = 2k+5, 2k+6, 2k+7$.

Such an assignment guarantees that the difference between any two channels that are in different groups is at least k. Every channel X is surrounded by channels that are in the other two groups, so there are no adjacent channels in the same group as X. Therefore, all channels surrounding X have labels that differ from X's label by at least k. In addition, no two cells with the same label are closer together than $4s$. These two properties establish this arrangement as a solution. The highest label used is $2k + 7$, so the bandwidth is $2k + 7$. This establishes a new upper limit on the span. $\square$

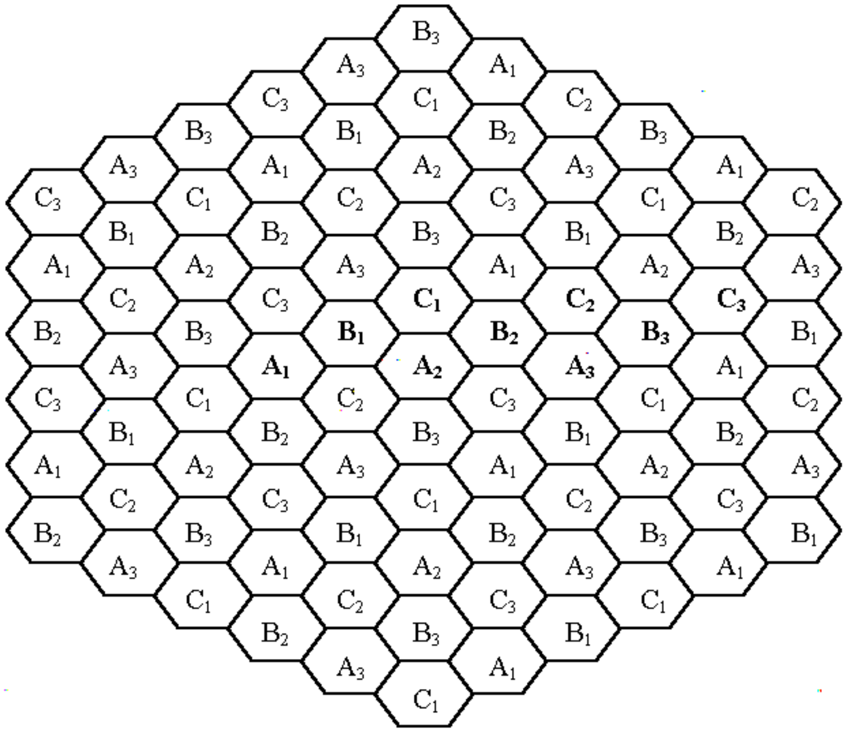


Figure 6. Design of a solution.

So, for $k \geq 4$, the span is either $2k+5$, $2k+6$, or $2k+7$. We need more building blocks to help us:

Lemma 1. *If there is no solution with bandwidth b, than the span is greater than b.*

Proof: Assume that there is a solution with bandwidth $m < b$. Replace the largest channel used with b. Since the relative spacing between labels either stays the same or grows, this also must be a solution, with bandwidth b. □

Lemma 2. *For any $k \geq 4$, given 6 consecutive labels, any selection of 5 of these cannot all be adjacent to a given cell X.*

Proof: Five hexagons all adjacent to a given hexagon include four edges between pairs (**Figure 7**).

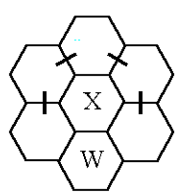


Figure 7. Situation of **Lemma 2**.

But for the six labels y, $y+1$, $y+2$, $y+3$, $y+4$, and $y+5$, there are at most three ways in which they can be adjacent (y with $y+4$, y with $y+5$, or $y+1$ with $y+5$, for $k=4$). Since no label can be used more than once, there is no way to label all five hexagons. $\square$

Lemma 3. *Any 4 hexagons that are all adjacent to a given hexagon must include at least 2 whose labels differ by more than k.*

Proof: Assume that there are 4 hexagons around a given hexagon such that the lowest label of the 4 is x and the highest label is $y \le x+k$. At least 2 edges are shared among the four hexagons (**Figure 8**). But there is at most one way in which the labels between x and y can be adjacent: x with $x+k$. Since no labels can be used more than once, there is no way to label all 4 hexagons. $\square$

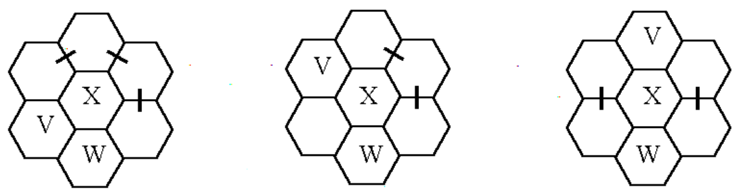


Figure 8. Situation of **Lemma 3**.

Theorem. *For $k \ge 4$, the span is at least $2k+7$.*

Proof: Suppose that the span is $2k+6$ and the label $k+1$ is used. The possible labels adjacent to $k+1$ are $1, 2k+1, \ldots, 2k+6$, which include 6 consecutive channels. But a selection of 5 out of 6 consecutive labels cannot all be adjacent to $k+1$ (**Lemma 2**). So any possible solution uses neither $k+1$ nor, by the contrapositive of the symmetry argument, $k+6$.

Next, suppose that the label $k+2$ is used. The possible labels adjacent to $k+2$ are $1, 2, 2k+2, \ldots, 2k+6$, which include 5 consecutive channels. But a selection of 4 out of 5 consecutive labels cannot all be adjacent to $k+2$ (**Lemma 3**). So any solution uses neither $k+2$ nor $k+5$ (by symmetry).

There are now 3 groups of labels remaining: $\{1, \ldots, k-1\}$ (group A), $\{k+3, k+4\}$ (group B), $\{k+8, \ldots, 2k+6\}$ (group C). No two channels in the same group can be adjacent, since they differ by less than k.

Suppose that a label X from group A is used. The possible adjacent labels must be from groups B and C. Since group B only has 2 channels, there must be at least 4 channels from C adjacent to X. But by **Lemma 3**, this is impossible; so no label from group A is used. By symmetry, no label from group C is used.

This leaves only the 2 channels in group B. Since every solution includes at least 7 distinct channels, there is no solution with a bandwidth of $2k+6$, so by **Lemma 1** the span is at least $2k+7$. $\square$

We have specified the span for all k: 7 for $k=1$, 9 for $k=2$, 12 for $k=3$, and $2k+7$ for $k \ge 4$.

Evaluation of the Model

Our model finds the span exactly for the constraints in requirements A, B, and C. This model still is limited by the assumptions of regular placement of transmitters, equal coverage around transmitters, and equal channel size. These constraints, however, may be too limiting or erroneous in real-world applications. We therefore examine three generalizations of the model.

Multiple Levels of Interference

Typically, in wireless communication the strength of a signal from a transmitter declines with the distance between receiver and transmitter. So interference caused by transmitters occupying the same or close frequencies also decreases with distance. In our main model, transmitters on the same frequency must be at least $4s$ apart, while transmitters on close frequencies must be $2s$ apart. We consider a generalization.

The distance between any two cells is 1 more than the minimum number of cells you must go through to go from the first cell to the second (a distance of 1 corresponds to $2s$). We consider a model where the amount of interference decreases linearly with distance. The needed separation f between any two cells as a function of distance d is $f = k(n-d)$, for $d \leq n$.

Theorem. *The span for* $n = 1$ *is* $2k+1$*.*

Proof: For any solution, there is a smallest label z that is used. Select two cells that are adjacent to this cell and adjacent to each other; these two cells must have different labels, x, y, with $x < y$. We know that $x \geq k + z$. But then $y \geq x + k$, so $y \geq 2k + z$. Since $z \geq 1$, we have $y \geq 2k+1$. This proves that the span is at least $2k+1$.

By setting $z = 1$, $x = k+1$, and $y = 2k+1$, and arranging them as indicated in **Figure 9**, we achieve $2k+1$ as the bandwidth, so the span is $2k+1$. □

Theorem. *For* $k = 1$ *and* $n \geq 2$ *, the span is no more than* $n^3 + n^2 - n + 1$*.*

Proof: We demonstrate our construction first for the case $n = 3$ in **Figure 10**. The general construction is shown in **Figure 11**, where $X, a, b = 1 + (a-1)(n-1) + (b-1)n^2$. The largest channel is $X, (n+1), (n+1) = 1 + (n+1-1)(n-1) + (n+1-1)n^2 = n^3 + n^2 - n + 1$. □

Theorem. *The span for any* k *is at least*

$$\begin{cases} 3k\left(\dfrac{n^2}{4} + \dfrac{n}{2}\right), & \text{for even } n; \\[2ex] \dfrac{3k(n^2-1)}{4} + 1, & \text{for odd } n. \end{cases}$$

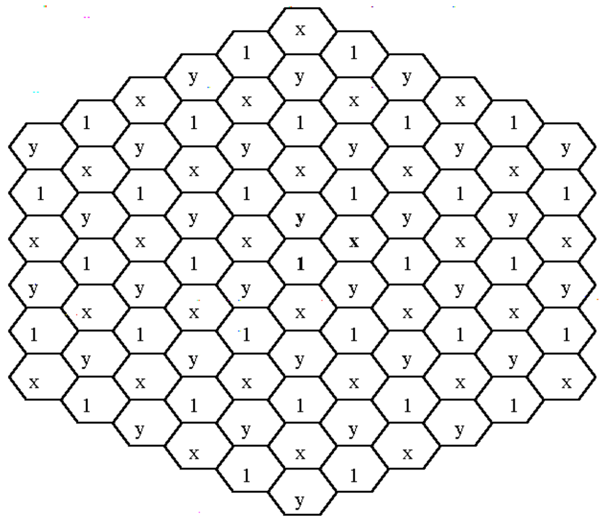


Figure 9. Solution with span $2k+1$.

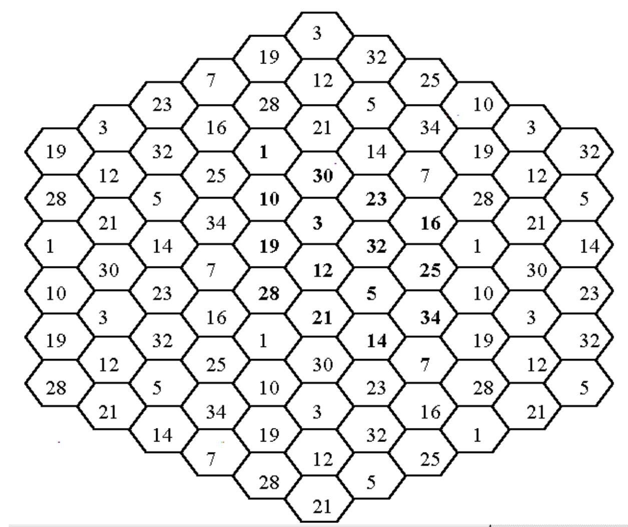


Figure 10. Construction for the case $n=3$.

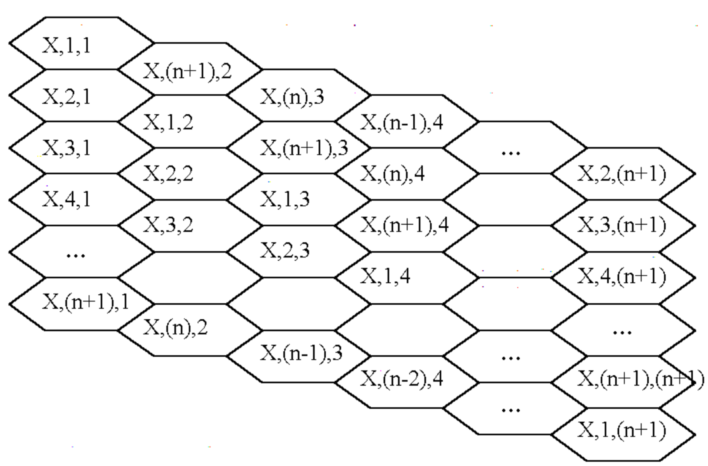


Figure 11. Construction for general n.

Proof: Case n even: Consider an arbitrary cell and the first $n/2$ rings of cells around it; all of the cells in this region are within a distance n of each other. Therefore, they must all be distinct, and they all must differ from each other by at least k. There are

$$3\left(\frac{n^2}{4}+\frac{n}{2}\right)+1$$

cells in this region, so the bandwidth must be at least as large.

Case n odd: Any solution for n is also a solution for $n-1$. Therefore the same formula applies with $n-1$ in place of n, yielding the value stated. □

Theorem. *Given a solution for some n and with bandwidth b, a solution exists for n and any k with bandwidth $kb-k+1$.*

Proof: Multiply all labels by k; this increases the separation between any two cells by a multiple of k. Subtract $k-1$ from all channels. This returns the lowest label to 1 and does not affect separation. □

We have established that the lower bound of the span for this modified model increases linearly with k and with n^2; we have established an upper bound that increases linearly with k and with n^3.

Weaknesses

This modification still assumes uniform arrangement of the transmitters. In addition, there is a discrepancy error between distance as we have defined it for hexagons and actual linear distance; however, an increase in hexagonal distance occurs if and only if there is an increase in actual distance as long as the hexagonal distance is less than 7.

Freedom of Transmitter Placement

We consider how far from the center of its cell a transmitter may be placed without violating any of the original constraints.

We assume that all transmitters can be displaced from the centers of their cells by an equal amount. The two interference constraints still apply: Transmitters within $2s$ of each other must still have a separation of at least k, and transmitters within $4s$ of each other must have a separation of at least 1. We consider two questions:

- What is the maximum freedom of displacement that can be allotted to transmitters before a solution developed by using our main model ceases to be a solution?
- What is the maximum freedom that can be allotted before a solution ceases to be a minimum solution?

Theorem. *If transmitters are displaced by less than $0.29s$, a solution developed by assuming that they are in the centers of the cells is still a solution.*

Proof: In the uniform arrangement of transmitters, the nearest transmitter that is more than $2s$ away from a given transmitter is at a distance of $s\sqrt{7} \approx 2.64s$. Their channel separation could potentially be less than 2, so we assume that it is. If both are moved toward each other by equal amounts, the $2s$ rule is violated when each has moved $0.32s$.

The nearest transmitter that is more than $4s$ away is at a distance of $s\sqrt{21}$. The two could have the same label. If they both move toward each other by equal amounts, the $4s$ rule is violated when they have both moved $0.29s$.

Therefore, as long all transmitters move less than $0.29s$, neither rule will be violated. $\square$

Theorem. *If transmitters are displaced by less than $0.13s$, a minimum solution developed by assuming that they are in the centers of the cells is still a minimum.*

Proof: In the uniform arrangement of transmitters, the farthest transmitter that less than $2s$ away from a given transmitter is at a distance of $s\sqrt{3} \approx 1.73s$. If both are moved away from each other by equal amounts, they are more than $2s$ apart when they have each moved $0.13s$.

The farthest transmitter that is less than $4s$ away is at a distance of $2s\sqrt{3} \approx 3.46s$. If they both move away from each other equal amounts, they will become more than $4s$ apart when they have both moved $0.27s$.

Therefore, as long all transmitters move less than $0.13s$, none of the constraints is affected. $\square$

Weaknesses

If transmitters are displaced from the centers, regions of their cells are more than s away from the transmitter. In addition, this approach does not consider the relative strength of signals as a function of the relative distance to the transmitters.

Rectilinear Constraints on Transmitter Placement

Our main model assumes that transmitters are arranged in a honeycomb pattern. However, due to city streets and township and county lines, a cellular service provider may be constrained to arrange transmitters rectilinearly. Therefore, we attempt to find the span under the original $2s$ and $4s$ constraints when the infinite plane is tiled with squares, each containing a transmitter in the center.

So, we consider a generalization of the $k = 2$ case (as specified in Requirements A and B) with an infinite grid of squares instead of hexagons. This requires a new definition of s, which now becomes the distance from the center of a square to one of its corners, so that s is still the maximum distance between a transmitter and a point in its cell. The cells considered within $2s$ and $4s$ from any given cell are shown in **Figure 12**. The inner four unmarked squares must have a separation of at least two from the center square. The outer ring of barred squares must not have the same label as the center (representing a separation of at least 2). The cells with a distance of exactly $2s$ and $4s$ are not included in the corresponding regions. We make this choice because when the cells are hexagons, each cell contains regions whose points are simultaneously within a distance of s of two transmitters; with square cells, there is no region of nonzero area where a point in one cell is within s of the transmitter in a diagonally adjacent cell.

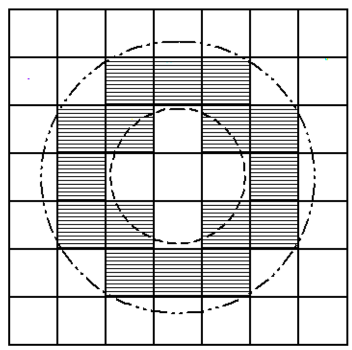


Figure 12. Square lattice of transmitters, with radii of $2s$ and $4s$ indicated.

Theorem. *The span is at least 8.*

Proof: In **Figure 13**, by the $4s$ condition, all of the boldfaced cells must be labeled distinctly. Therefore, there must be at least 7 distinct labels. Assume that there is a solution with a bandwidth of 7. Then all of the cells that are not in boldface must be labeled as indicated. However, all 7 of the labels are within $4s$ of the cell indicated with a question mark. Therefore, 7 labels are insufficient. □

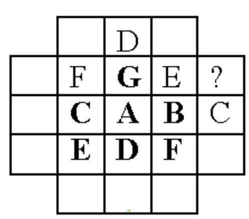


Figure 13. Seven labels are not enough . . .

Theorem. *The span is 8.*

Proof: The construction of a solution with a bandwidth of 8 is shown in **Figure 14**. □

7	4	1	6	3	8	5	2	7	4	1	6	3	8
3	8	5	2	7	4	1	6	3	8	5	2	7	4
5	2	7	4	1	6	3	8	5	2	7	4	1	6
1	6	3	8	5	2	7	4	1	6	3	8	5	2
7	4	1	6	3	8	5	2	7	4	1	6	3	8
3	8	5	2	7	4	1	6	3	8	5	2	7	4
5	2	7	4	1	6	3	8	5	2	7	4	1	6
1	6	3	8	5	2	7	4	1	6	3	8	5	2

Figure 14. . . . but eight are.

The block of repeating channels in **Figure 14** is reflected about a horizontal axis and then tiled to the right and left. In this fashion a solution for an infinite array of squares can be constructed.

Weaknesses

This modification still assumes that transmitters are placed regularly, that channels with a separation of more than 2 never interfere with each other, and that there is no interference beyond a radius of $4s$.

News Release

Let's say that you work for a company that provides cellular telephone service. Your job is to decide how the company should go about expanding into a new area. You decide where transmission towers will be placed and what channels will be needed.

You know that gaps in coverage must be avoided. There must not be any places in the new service area where a customer's cell telephone reads "We're sorry, you're outside the coverage area."

One way to ensure that there are no "holes" or gaps in coverage is simply to put up lots and lots of towers. "Put up a tower every half-mile," you tell your boss. Your boss says that the idea is interesting, and would certainly guarantee complete coverage, but that another essential goal is to minimize the number of towers needed. Each tower costs tens of thousands of dollars.

So you return to your cubicle and do some research. You learn that the signals from a transmission tower get weaker as a customer's phone gets farther and farther from the tower. You find out that there is a distance, say, five miles, beyond which clear reception can no longer be guaranteed. So you take a map of the region and start drawing circles whose radius is this distance (reduced to the scale of the map, of course). Right away you realize that the circles will have to overlap a bit. There are some quarters sitting on your desk, and you quickly see that one quarter can be surrounded by six other quarters, so that all the quarters are pushed together as close as possible. But there are little triangular holes in between them, so you have to overlap them a bit so that the area that they cover has no holes.

This, in fact, turns out to be the most efficient way to cover a plane with circles—there is as little overlap as possible. So you cover your map with circles in this fashion and decide to place a transmission tower in the center of each circle. Since the two neighboring circles overlap a bit, to eliminate those pesky triangular holes you split the overlap area between them. Since every circle is surrounded by six other circles, you do this for all six overlapping areas. Now the region (cell) that is serviced by any transmission tower is a hexagon. In fact, the hexagons form a honeycomb pattern on your map, and you wonder why you didn't think of that before. The distance from the center to the corner of any hexagon is as far as you can get from the center without leaving the hexagon, so this distance must be the five miles that you determined earlier is as far as a customer can get from a tower and still be assured of coverage.

Quite excited with your discovery, you run to your boss shouting, "Eureka! Hexagons!" Your boss smiles, a bit condescendingly, and informs you that they already knew that hexagons were the best way to cover a large area. What your boss really needs you to determine is how many channels the company will have to purchase from the FCC in order to have enough to for all the towers. "That's easy," you respond, still not realizing that nothing is as easy as it seems, "just buy one channel for each tower!" Your boss sighs, silently wondering if you are the right person for this job. Then your boss reminds you that not spending large amounts of unnecessary money is also an essential part of doing business. The company needs to know: What is the smallest number of channels that they will have to buy? And how should the channels be assigned to the towers?

Having finally learned to be cautious, you suppress the urge to blurt out, "Just put the same channel on all the towers!" Instead, you begin asking questions. Your manager refers you to the engineering department. This is where the problem gets interesting.

First, you are told that if a transmitter in some cell uses a particular channel, then none of the cells in the first ring of six hexagons around it can share that channel. This is because when you move from one cell to another, the way that the cell phone stops talking to one tower and starts talking to the tower in the new cell is by changing channels. If two cells right next to each other used the same channel, a phone in between would not know which one to listen to!

Next, you learn that none of the cells in second ring of 12 hexagons can use the same channel either. In fact, the closest that two cells that use the same channel can be is three hexagons apart. So the same channel can be used over and over again, but the cells that use it must be spread out so that they are all at least three cells apart from each other.

Finally, you find out that cells that are right next to each other cannot use channels that are too close together in frequency. When you ask why, the engineers begin enthusiastically telling you about something called "spectral spreading." You decide that it is better to not know why it is true, only that neighboring cells must use channels that are separated by some number of channels. For instance, suppose neighboring channels must be separated by at least four. Then if one cell uses channel 10, then perhaps all its neighbors must use channels that are at least four away, that is, none of the neighbors can use 7, 8, 9, 11, 12, or 13. You are told that for this new region, it has not been determined yet just how far apart the channels in neighboring cells must be from each other. Perhaps it depends on the expected call volume and on atmospheric conditions in the new area. But as soon as this channel separation is determined, you will be expected to say immediately how many channels are needed and how they should be distributed among the towers.

When you return to your desk, you find a message on your chair. It seems that your boss neglected to mention an important detail: It does not matter how many channels the company actually uses, but what is crucial is *how big a block of consecutive channels* in the frequency spectrum the company occupies. For example, if you use only channels 11 and 20, you still have to pay for the block of ten channels between 11 and 20.

This is the problem that our mathematical modeling team recently solved. We determined that if neighboring cells must be at least two channels apart, then a region of any size can be covered with a block of 9 consecutive channels. If a separation of three channels is required, any region can be covered with 12 channels. If a separation of four or more channels is required between neighboring cell, then the necessary number of channels can be found by doubling the separation and adding 7. For example, if neighboring cells must be 5 channels apart, then 17 channels must be purchased from the FCC. We also determined that these are the best possible solutions. That is, there is no way to cover a large region with fewer channels, without breaking some of the rules that the engineering department requires.

But what is interesting about our result is not the number of channels required but how the channels are distributed among the towers. An example will illustrate this point. Suppose again that neighboring cells must be five channels apart; then our model calls for 17 channels to be purchased, say channels 1 through 17. But the only channels that our solution uses are 1, 2, 3, 8, 9, 10, 15, 16, and 17! We spread out the first three channels among the cells so that none of them are adjacent to each other. This covers one-third of the all the cells. We repeat this with channels 8, 9, and 10. Now two-thirds are covered. The rest are covered with 15, 16, and 17. We proved that this is the best possible solution. It is surprising that the best solution has 8 channels in the middle of its block that are not even used, but this is indeed the case. What the unused channels do is divide the remaining channels into three groups. The groups are far enough apart that any channel in one group can be surrounded by channels in the other two groups.

Utilize the Limited Frequency Resources Efficiently

Chu Rui
Xiu Baoxin
Zong Ruidi
National University of Defence Technology
Chang Sha, Hunan, China

Advisor: Wu Meng Da

Introduction

For Requirements A, B, and C, we find an optimal assignment. We find efficient strategies for assignment and the spans for $k = 2$ (9), $k = 3$ (12), and $k \geq 4$ ($2k + 7$) for the case of two levels of interference.

For Requirement D, to check all possible assignments is impractical. Instead, we devise a heuristic algorithm to produce a near-optimal assignment and span.

We also consider other important factors, such as cell-splitting and duopoly.

Analysis of the Problem

We first define

- **successful assignment**: An assignment of channels that satisfies all constraints.
- **span**: The minimum, over all successful assignments, of the largest channel used.

Our goal is to find the span and a successful assignment under given constraints:

- **Constraint 1:** The channels of transmitters within distance $2s$ differ by at least k.
- **Constraint 2:** No transmitters within $4s$ of each other can use the same channel.

The UMAP Journal 21 (3) (2000) 343–356.

Assumptions

- The region is partitioned into a grid of regular hexagons (called *cells*), and each transmitter is located at the center of a hexagon.
- The frequency spectrum is divided into regularly spaced channels represented by positive integers.
- A channel can be used by more than one transmitter, provided interference is avoided.
- Each transmitter is assigned a single channel.
- Each transmitter covers its entire cell, and the effect of landform can be ignored.

Table 1.
Notation used.

s	length of a side of a hexagon
$A, B, \ldots$	cells or transmitters
$a, b, \ldots$	channels
$\omega_i s$	distance constraint of the ith level of interference
k_i	frequency constraint of the ith level of interference
k	frequency constraint of first level of interference in Requirement C
p, q	shift parameters
$d(A, B)$	distance between cells A and B

Model Design and Results

Requirement A

We find an optimal result via a recursive backtracking algorithm. Our goal is to see if we can use the integers up through n to satisfy the constraints. We loop from $n = 1$ to $n =$ number of cells (giving every cell a distinct channel must work). We order the cells and try each in turn to see if it can be numbered with an integer between 1 and n: if so, we proceed to the next cell; if not, we renumber the previous cell. If all cells can be numbered, then we have a successful assignment with n channels; otherwise, n channels are not enough.

Proposition 1. *For requirement A, the span is 9.*

Proof: Every cell is adjacent to six others, and these cells are all within $4s$ distance of each other; so according to **Constraint 2**, these seven cells must be assigned different channels. Furthermore, if one cell has channel m, the six adjacent cells cannot have $m + 1$ or $m - 1$, according to **Constraint 1**. Hence, the span must be at least 9. Our algorithm finds a successful assignment with 9 channels (**Figure 1**), so 9 is the span. □

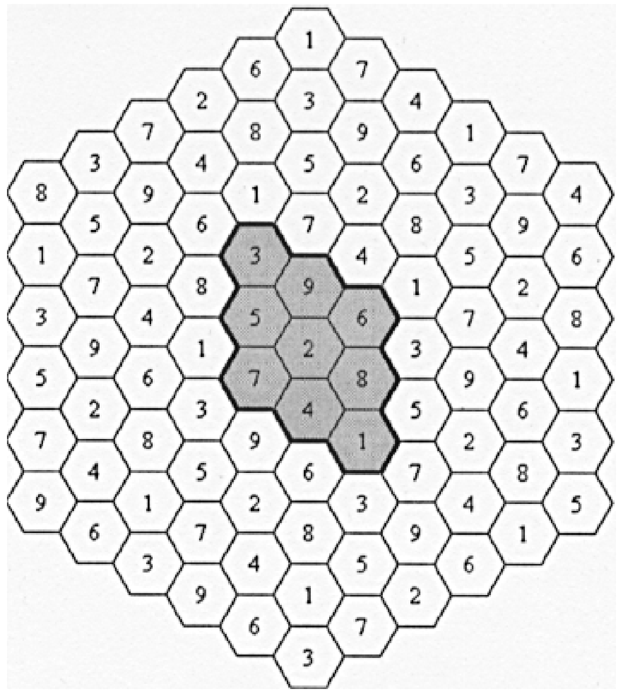


Figure 1. For $k = 2$, the span is 9.

Requirement B

The shaded part in **Figure 1** can be expanded arbitrarily far in all directions, so for Requirement B the span is still 9.

Requirement C

For Requirement C, there are still two levels of interference, but the k of **Constraint 1** is allowed to vary.

Our algorithm finds the successful assignments of **Figure 2** for $k = 3, 4, 5$, for which the given spans can be proved easily.

The shaded parts of **Figure 2** show patterns that allow the same radio frequencies to be reused throughout. We systematize the reuse pattern. We begin with a pair of integers $p \geq q$, which we call *shift parameters*. In a hexagonal tiling there are six "chains" of hexagons emanating in different directions from each hexagon. Starting at any cell, we proceed as follows: Move p cells along any chain, turn CCW 60°, then move q cells along the chain in that direction. The original cell and the qth cells so located in each direction are *co-channel* cells (**Figure 3**).

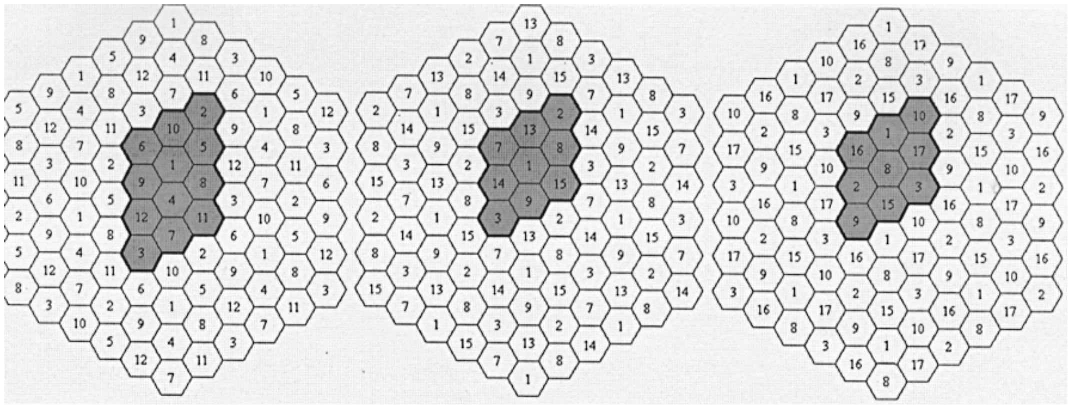


Figure 2. **a.** $k = 3$: the span is 12. **b.** $k = 4$: the span is 15. **Figure c.** $k = 5$: the span is 17.

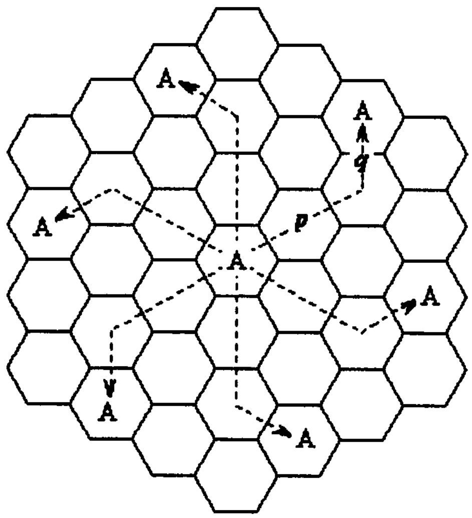


Figure 3. A cell and co-channel cells.

We repeat the procedure for a different starting cell until all cells in the region are assigned. The region can then be divided into a *cluster* of cells, such that transmitters in the same cluster are assigned different channels. The form of the cluster is determined by the shift parameters p and q, and it can be proved that the number of cells in a cluster is $p^2 + pq + q^2$.

In particular, we want to find a suitable cluster of cells for all $k \geq 4$. To minimize the width of the interval of the frequency spectrum, we adopt the cluster of 9 cells in **Figure 2c**; it can be proved that $2k + 7$ is the best result for the span. This reuse pattern is shown in **Figure 4**.

We define the three integer sets N_1, N_2, N_3:

$$N_1 = \{1, 2, \ldots, k\},\ N_2 = \{k+1, k+2, \ldots, 2k\},\ N_3 = \{2k+1, 2k+2, \ldots, 2k+6\}.$$

Proposition 2. *In a successful assignment with largest channel no more than* $2k + 6$ $(k \geq 6)$, *channels for two transmitters at a distance of* $3s$ *must belong to the same set* N_i.

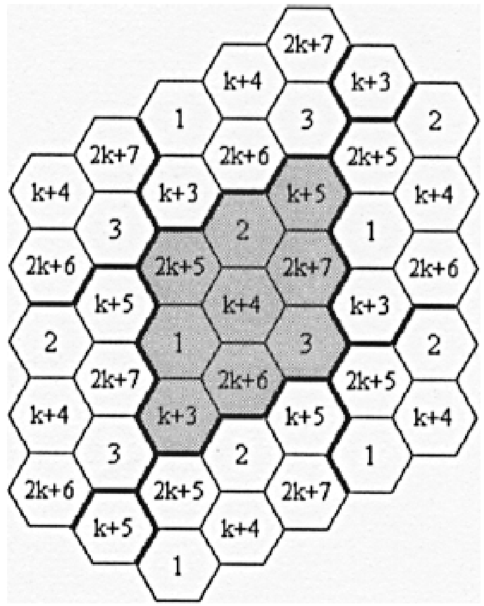


Figure 4. Cell reuse pattern using 9 channels.

Proof: Let A, B be the transmitters, and let C, D be transmitters adjacent to both (**Figure 5**); a, b, c, d are the corresponding channels in a successful assignment. Considering **Constraint 1**, b, c, d must belong to three different sets N_1, N_2, N_3; similarly, a, c, d must belong to three different sets. Hence a, b must belong to the same set. □

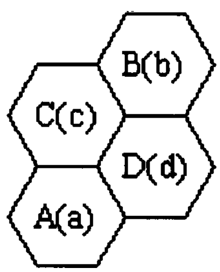


Figure 5. Situation of **Proposition 2**.

Proposition 3. *The span is* $2k + 6$ *for* $k \geq 6$.

Proof: Suppose not. Select any cell A not at the edge of the given region. If its channel a is in N_2, then the six channels of the cells adjacent to A must belong to N_1 or N_3. By **Proposition 2**, three of the six channels belong to N_1, and they are different from each other; hence, we deduce that $a \geq k + 3$ by **Constraint 1**.

From the shaded part of **Figure 6**, select any cell B whose channel b is in N_3. Let the channels for the cells adjacent to B be c, d, e, f, g, h. From **Proposition 2** and **Constraint 1**, we may assume that $d, f, h \in N_1$ and $c, e, g \in N_2$; thus, we have

$$\min\{c, e, g\} \geq k + 3 \longrightarrow \max\{c, e, g\} \geq k + 5 \longrightarrow b \geq \max\{c, e, g\} + k \geq 2k + 5.$$

So, in N_3 there are only two integers $(2k+5, 2k+6)$ that can be assigned to the shaded part of **Figure 6**, which is impossible.

Hence, the supposition that the span is less than or equal to $2k+6$ is not true. We already have a successful assignment with largest channel $2k+7$, so the span is $2k+7$. $\square$

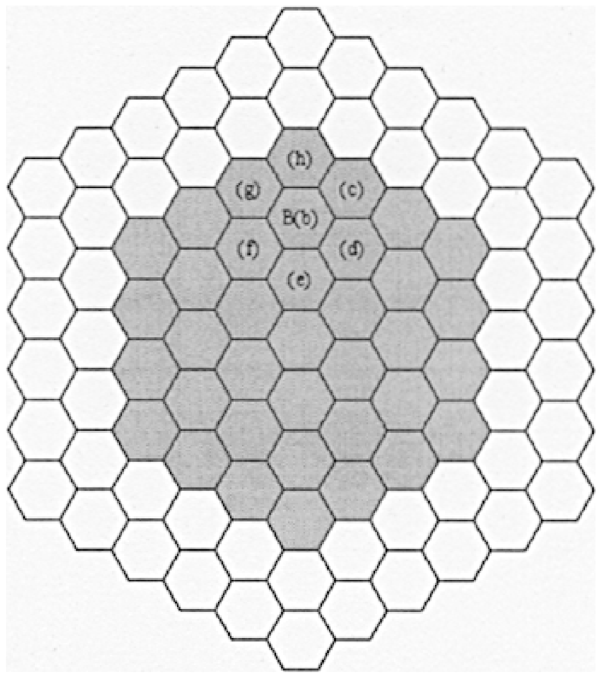


Figure 6. Situation of **Proposition 3**.

Requirement D

Several Levels of Interference

Our backtracking algorithm's workload increases rapidly with the number of interference levels, and we don't know what kind of clustered region could attain the span. So we turn to an approximate algorithm.

We notice from our results for Requirement C that

- There are several groups of channels; channels in the same group are close and evenly distributed, while channels in different groups differ greatly.
- In a single cluster, the channels are all different.

Considering these observations, we guess that the distance between any two adjacent channel-reuse cells should be constant. We use this as the basis for the

Heuristic Skip Algorithm (HSA)

Let there be n levels of interference, and let the channels for transmitters within $\omega_i s$ of each other have to differ by at least k_i. Let Ω be a "cell set" and let $\alpha \in \{1, \ldots, n\}$ be a control parameter.

Step 1. Choose the center cell A of the region as the initial cell:

$$j := \alpha, \qquad l := 1.$$

Step 2. If $j = 0$, stop;
else, pick out all cells A_i that have not been numbered but satisfy
$\omega_{j-1}s \leq d\,(A_i, A) \leq \omega_j s.$

Step 3. If there are not such A_i, then set $j := j - 1$ and go to Step 2;
else, choose one cell nearest to A from A_i and assign the minimal feasible channel to it. If $l = 1$, denote the selected cell by B and denote the shift parameters from A to B by p, q.

Step 4. Add the selected cell to Ω.

Step 5. Start with any cell in Ω as a reference, move p cells along any chain of hexagons, turn CCW 60°, and move q cells along the chain in the new direction. Assign the minimal feasible channel to this new cell and add it to Ω.

Step 6. If there is no starting cell in Step 5, set $\Omega := \phi$ and $l := l + 1$ and return to Step 2;
else, repeat Step 5.

When the algorithm ends, every cell is numbered. The control parameter α determines the distance between any two adjacent channel-reuse cells, and we can execute the algorithm repeatedly with different values of α to get the best result.

Results

- 2 levels of interference (e.g., $\omega_1 = 2, \omega_2 = 4, k_1 = 5, k_2 = 1$): The largest channel assigned by HSA is 17, which agrees with the optimal result earlier.

- 3 levels of interference (e.g., $\omega_1 = 2, \omega_2 = 4, \omega_3 = 6, k_1 = 5, k_2 = 3, k_3 = 1$): The largest channel assigned by HSA is 33; the cluster pattern is shown in **Figure 7a**.

- 4 levels of interference (e.g., $\omega_1 = 2, \omega_2 = 3, \omega_3 = 4, \omega_4 = 6, k_1 = 4, k_2 = 3, k_3 = 2, k_4 = 1$): The largest channel assigned by HSA is 29; the cluster pattern is shown in **Figured 7b**.

Irregular Transmitter Placement

Let r be the largest distance between a transmitter and the center of its hexagon. For $r \leq 0.134s$ and two levels of interference, the results apply as before, and a similar analysis can be made for other cases and numbers of levels of interference.

For larger r, we still use HSA. But first, since the position of transmitters is irregular, some might be missed when we use HSA to assign channels. We

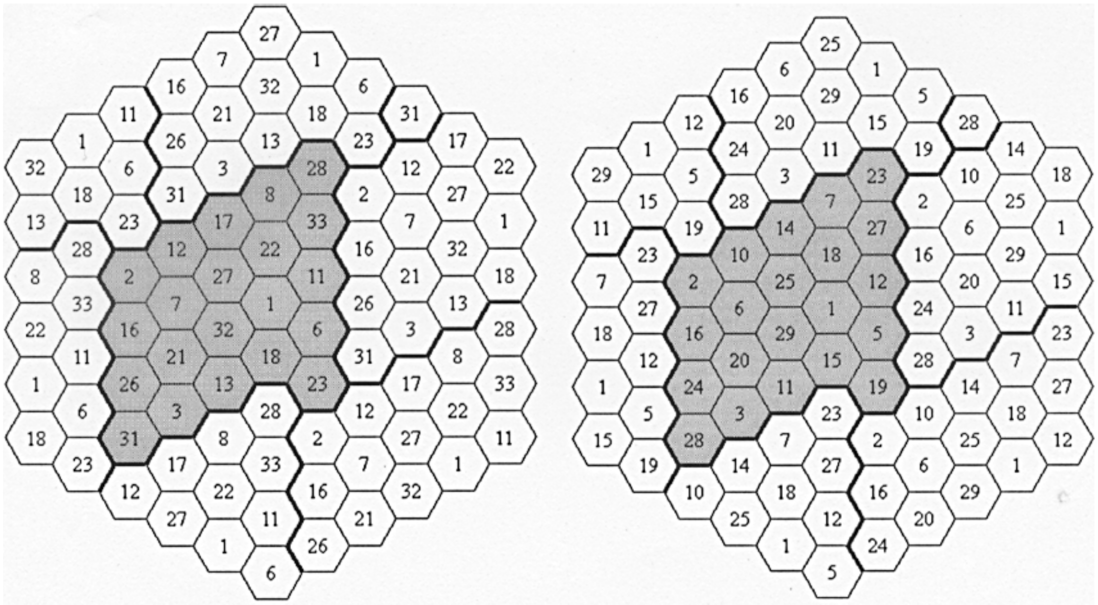


Figure 7. a. Assignment for 3 levels of interference. **b.** Assignment for 4 levels of interference.

overcome this problem by improving HSA so that when it determines which transmitter is to be assigned a channel, it ignores shift parameters, but it does consider the position of the transmitter when assigning a channel.

Result

- 2 levels of interference (e.g., $\omega_1 = 2, \omega_2 = 4, k_1 = 4, k_2 = 1$): To simulate reality, we randomly choose 80% of the transmitters move by $0.134s$ and the others to move by $0.3s$. The largest channel is 16, compared with 15 for regular placement.

Analysis of Results

We use HSA to solve the problems under various conditions. Attenuation of radio signals follows a log normal distribution. Since the radius of a cell is several kilometers, we solve only the problem of $6s$ interference, which should suffice in reality.

Two or More Levels

In Requirements A, B, and C, only two levels of interference are taken into account, with $\omega_1 = 2$ and $\omega_2 = 4$. For $k = 2$, HSA gives 11 channels, while the span is 9; for $k = 3$, HSA gives 13, while the span is 12; for $k = 4, \ldots, 10$, HSA gives the span.

For Requirement D, with three levels of interference ($\omega_1 = 2, \omega_2 = 3, \omega_3 = 4$), HSA gives a largest channel of $3k_1 + 6$, which we feel is very close to the span. **Figure 8** shows the frequency reuse pattern of a cluster of 12 cells.

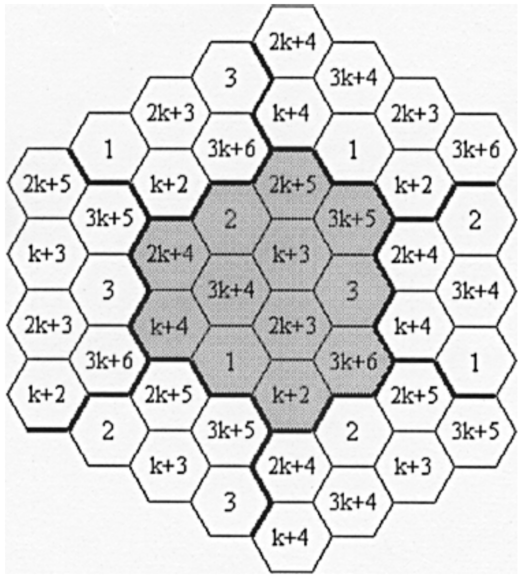


Figure 8. Cellular reuse pattern using a cluster of 12 channels.

Table 1 gives results for various combinations of the parameters for the case of 3 levels of interference; **Table 2** gives results for 4 levels.

For three levels and k_2 small compared to k_1, and for four levels with $k_3 = 2$, the HSA results are smaller than for other cases, because the algorithm can adopt a more rational cluster structure and assign channels more economically. This fact indicates that the result of HSA is determined not by one or two parameters but by all parameters and that HSA makes full use of the information of the constraints; hence, HSA may give a comparatively good result.

Irregular Transmitter Placement

Tables 3–4 give the results, which varies only slightly when the proportion of transmitters moved more than $0.3s$ is less than 10%.

Table 1.

Results for 3 levels ($\omega_1 = 2, \omega_2 = 4, \omega_3 = 6$).

k_1	k_2	k_3	HSA
3	2	1	27
4	2	1	27
	3	1	33
5	2	1	33
	3	1	33
	4	1	39
6	2	1	38
	3	1	39
	4	1	39
	5	1	45
7	2	1	38
	3	1	45
	4	1	45
	5	1	45
	6	1	51
8	2	1	40
	3	1	51
	4	1	51
	5	1	51
	6	1	51
	7	1	57
9	2	1	42
	3	1	53
	4	1	57
	5	1	57
	6	1	57
	7	1	57
	8	1	63

Table 2.
Results for 4 levels ($\omega_1 = 2, \omega_2 = 3, \omega_3 = 4, \omega_4 = 6$).

k_1	k_2	k_3	k_4	HSA
4	3	2	1	29
5	3	2	1	32
	4	2	1	34
		3	1	35
6	3	2	1	37
	4	2	1	37
		3	1	39
	5	2	1	37
		3	1	41
		4	1	41
7	3	2	1	38
	4	2	1	38
		3	1	45
	5	2	1	38
		3	1	45
		4	1	45
	6	2	1	40
		3	1	47
		4	1	47
		5	1	47

Table 3.
Irregular transmitter placement, 2 levels of interference ($\omega_1 = 2, \omega_2 = 4$).

% moved $> 0.3s$	$k_1 = 4, k_2 = 1$	$k_1 = 6, k_2 = 1$
0	15	19
5	15	19
10	15	19
20	16	27
25	20	29
30	21	33
35	21	31
40	21	31
45	15	29
50	21	26

Table 4.
Irregular transmitter placement, 3 levels of interference ($\omega_1 = 2, \omega_2 = 4, \omega_3 = 6$).

% moved > $0.3s$	$k_1 = 4, k_2 = 3, k + 3 = 1$	$k_1 = 6, k_2 = 4, k_3 = 1$
0	33	50
5	33	41
10	33	42
20	38	44
25	38	46
30	38	46
35	34	43
40	34	54
45	36	51
50	39	57

Further Discussion

Cell Splitting

One advantage of cellular service is its ability to keep up with changing customer demands. If the customer base approaches full capacity, key cells can be divided into a number of smaller cells, each broadcasting at lower power, and channels can be reassigned to increase the volume of customers (**Figure 9**).

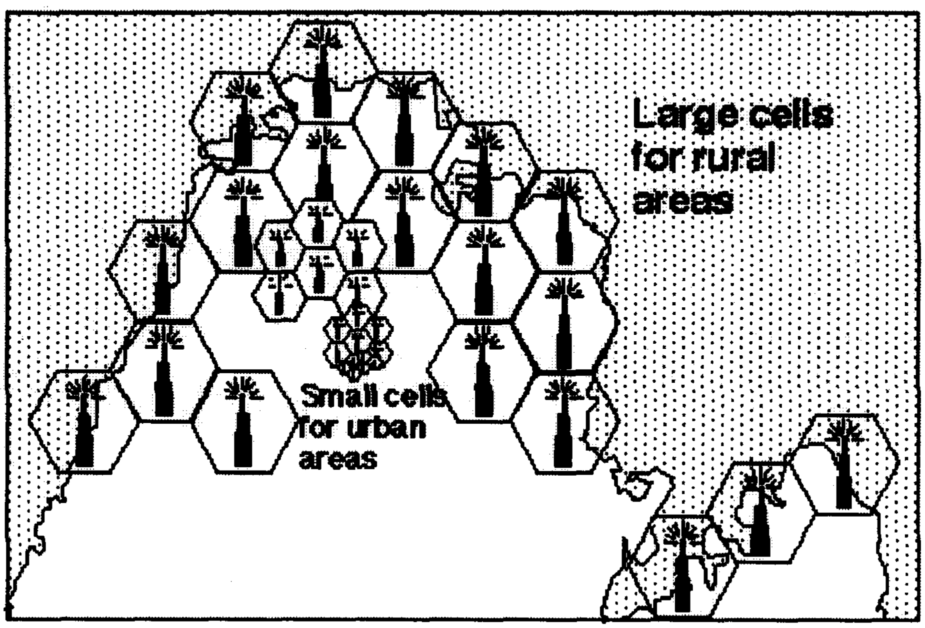


Figure 9. Cell-splitting.

Assumptions

- The radius of a new cell is half that of an old one, and the power of the new transmitter is half that of the old one; as a result, the distance constraints between new cells become half that of the old.
- The new cells plus unsplit old ones cover the entire region.
- The number of levels of interference between any two old cells are unchanged.
- The number of levels of interference between an old cell and a new cell are the same as between the old cells.

Strategies

- Strategy 1: Assign the new cells first, then the old ones.
- Strategy 2: Assign the old cells first, then the new ones.

Results

We allot channels by HSA in the area with split shells that is shown in **Figure 10**. For two levels of interference ($\omega_1 = 2, \omega_2 = 4, k_1 = 4, k_2 = 1$), Strategy 1 uses 28 channels and Strategy 2 uses 27.

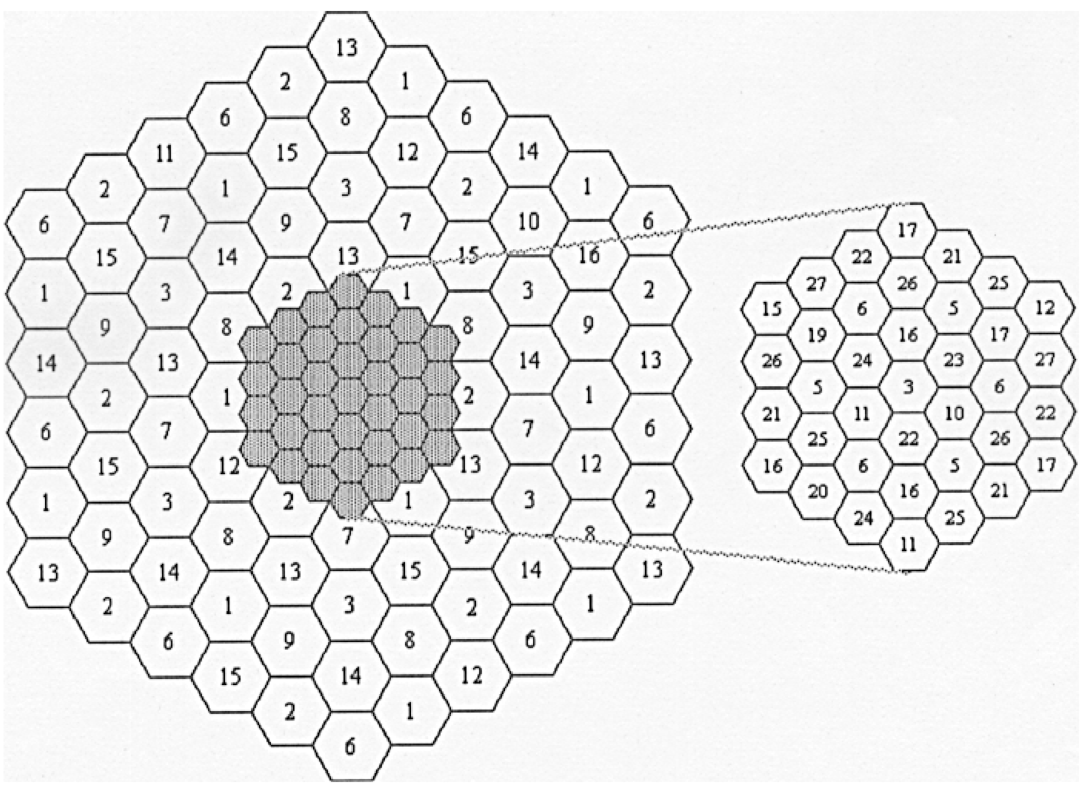


Figure 10. Test region for splitting strategies; result for Strategy 2 is shown (largest channel is 27).

Duopoly

If there are two providers in a region, we must assign their channels at the same time to avoid cross interference between the two systems.

Strengths

For two levels of interference, we determine the span as a function of k, prove optimality of the result, and give an efficient strategy for assigning channels.

The HSA algorithm is adaptable to large areas, irregular transmitter placement, and splitting cells. It seems to give results close to the span, together with a cluster of cells that allows simple assignment of channels.

The HSA algorithm is polynomial-bounded (in class $\mathcal{P}$); for the test cases examined, it gave results (on a PC) within 5 sec.

Weaknesses

The result of the HSA algorithm may not be optimal.

Although we conjecture that the span is $3k_1 + 6$ for the situation with parameters $\omega_1 = 2, \omega_2 = 3, \omega_3 = 4$, we cannot prove it.

References

EL208 Radio Communications Engineering: Week 6: Introduction to Cellular Telephones. `http://homepages.unl.ac.uk/~fusterg/el208/week6.htm`.

Elliott, Scott D., and Daniel J. Dailey. 1995. *Wireless Communications for Intelligent Transportation Systems.* Boston, MA: Artch House.

Simon, Haykin. 1981. *Communications Systems.* New York: Wiley.

Yee, William C.Y. 1982. *Mobile Communications Engineering.* New York: McGraw-Hill.

Groovin' with the Big Band(width)

Daniel J. Durand
Jacob M. Kline
Kevin M. Woods
Wake Forest University
Winston–Salem, NC

Advisor: Edward Allen

Introduction

We have a planar surface divided into hexagonal cells with sides of length s. In **Figure 1**, all of the "first concentric" (striped) cells lie within $2s$ of the central cell, whereas all of the dotted cells and all the "second concentric" (dotted plus striped) cells lie within $4s$ of the central cell.

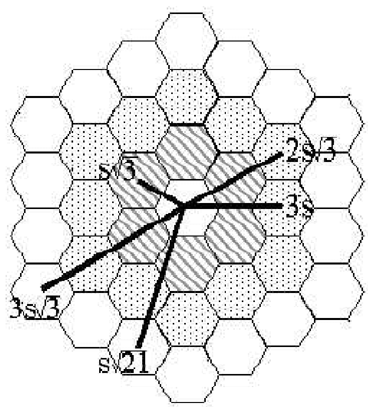


Figure 1. Diagram of interference regions.

The First Case

We analyze the case where transmitters within $4s$ must differ by at least 1 channel and those within $2s$ must differ by at least 2 channels. We show that the span of the network is 9 for both the finite grid in the problem statement and for the infinite plane; the answers to Requirements A and B are identical.

The UMAP Journal 21 (3) (2000) 357–367.

We begin by considering the "first concentric," shown in **Figure 1** by the central cell and the ring of striped cells surrounding it. Since any cell of the first concentric is within $4s$ of all the others, each cell in it must be assigned a distinct integer; so the span cannot be 7 or less.

A more careful examination reveals that the span cannot be 8. Suppose that it were. Consider three adjacent hexagons that share a common vertex, we find that only one cell can be assigned a 1 and only one cell can be assigned an 8. Thus, the remaining cell must be assigned some numbers between 2 and 7. Consider this last cell as the center of a first concentric and assign it n. Then the $2s$ constraint dictates that the ring of 6 cells surrounding it cannot be assigned numbers $n-1$, n, or $n+1$. Their assignments must also be distinct from one another, since all cells within the first concentric are within $4s$ of each other. To make these cell assignments, we need six integers other than $n-1$, n, or $n+1$, or at least 9 numbers altogether; so the span must be at least 9.

Figure 2 shows a solution with 9, so the span is 9. The central column in gray is the sequence 1, 3, 5, 7, 9, 2, 4, 6, 8 repeated over and over. The column to the right of it (dotted) is the same sequence but shifted down 3 cells; the striped column to the left of center is the same sequence shifted up 3 cells. Repeat this process of shifting up or down indefinitely to the left and right. Look at each 1 in the pattern (in black). The column to the left of each 1 is always shifted up by 3, and the column to the right is always shifted down by 3. Therefore, each 1 must have the same neighbors. The cells within $2s$ of the 1s differ from it by at least 2, and those within $4s$ by at least 1; so the pattern meets the constraints. Checking the neighbors of the other numbers 2 through 9 shows that they meet the constraints also. This pattern can fill the grid supplied in the problem, or it can be extended arbitrarily far left and right and also up and down to cover the plane. This pattern is unique, not including rotations and reflections. [EDITOR'S NOTE: We omit the proof of this fact.]

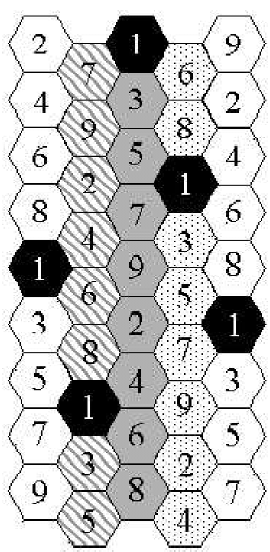


Figure 2. Solution with 9 channels.

Generalization: Differing k

In this section, we maintain the constraint that transmitters within a distance $4s$ of one another cannot use the same channel but generalize the second constraint, so that transmitters within a distance $2s$ of one another must have channels whose assignment numbers differ by k. The previous section treated the case $k = 2$.

We show that for $k = 1$, the span is 7; for $k = 3$, the span is 12; and for $k > 3$, the span is $2k + 7$.

First, we show that for all k, $2k + 5$ is a lower bound for the span. Suppose that we have a channel configuration that uses only 1 through $2k + 4$; this will lead to contradiction. Let $A = \{1, 2, \ldots, k\}$ and $B = \{k + 5, k + 2, \ldots, 2k + 4\}$. All numbers in A are within k of each other, as are all numbers in B. Consider three adjacent hexagons that share a common vertex. At most one of these three can be assigned an element of A and at most one can be assigned an element of B, so the third must be assigned some channel n between $k + 1$ and $k + 4$. Consider a first concentric in which the central cell has been assigned this integer n. The $2s$ constraint dictates that the 6 adjoining cells cannot be assigned numbers $n-k+1, n-k+2, \ldots, n+k-2, n+k-1$. Their assignments must also be distinct from one another, since all cells within the first concentric are within $4s$ of each other. To make these cell assignments, we need six integers other than $n - k + 1$ through $n + k - 1$. This means that we need $6 + (2k - 1) = 2k + 5$ integers. Therefore, we cannot make proper channel assignments using only the integers 1 through $2k + 4$.

$k = 1$

When $k = 1$, the $2s$ constraint is subordinate to the $4s$ constraint. The span must be at least $2k + 5 = 7$. In fact, we can complete the grid using a span of exactly 7. As in **Figure 2**, the central column is a sequence of numbers repeated over and over, in this case the sequence 1, 2, 3, 4, 5, 6, 7. Also as in **Figure 2**, the adjacent column on the right contains the same sequence shifted down 3 cells, and the adjacent column on the left contains the same sequence shifted up 3 cells. For example, the 1 in the column to the right is between the 3 and 4 of the central column. As in the $k = 2$ constraint, every occurrence of each integer would have identical neighbors. Using this pattern, we can construct a satisfactory network. Moreover, since we have proved that the span must be greater than 6, our construction demonstrates that the span is exactly 7.

$k = 3$

We show that no assignment exists that uses only 1 through 11 and provide an example that works for 12, thereby demonstrating that the span is 12.

Assertion A: *The span must be greater than 11.*

Proof of A: By contradiction. Suppose that the span is 11. We show that several channel numbers cannot appear and use these facts for our final contradiction.

Case A1: Suppose that some transmitter is assigned channel 3. Consider a first concentric about a central cell assigned channel 3. No transmitters in the first concentric can use channels 1, 2, 3, 4, or 5, because they are all within $2s$ of the center transmitter operating on 3. We are left with six viable channels, 6, 7, 8, 9, 10, and 11, all of which must be used to provide distinct assignments to the cells surrounding the center cell. Clearly, channel 8 must be used somewhere in the first concentric. We must then use two of the five remaining channels (6, 7, 9, 10, 11) in two empty cells of the first concentric lying to either side of 8. However, this is not possible, since placing either 6, 7, 9, or 10 in either of these cells would violate the $2s$ requirement (as 6, 7, 9, 10 are all within 3 of 8). It follows that no transmitter can be assigned channel 3.

Case A2: Suppose that some transmitter is assigned channel 9. Since channels n and m within a distance $2s$ of each other must differ by at least k, we have that $|n - m| \geq k$. If we flip all channel numbers m to $(\text{span} + 1 - m)$, then $|(\text{span} - n) - (\text{span} - m)| = |m - n| = |n - m| \geq k$. Thus, the set of new channels functions identically under the $2s$ and $4s$ constraints. The channel numbers remain between $12 - 11 = 1$ and $12 - 1 = 11$, so we have a correct channel assignment, and the span remains 11. So if some transmitter is assigned channel 9, a flip produces a configuration with a channel $12-9 = 3$, which Case A1 shows is impossible. Therefore, no transmitter can be assigned channel 9.

Case A3: Assume that some transmitter is assigned channel 10. Consider the first concentric around a channel 10 (in gray), as shown in **Figure 3**.

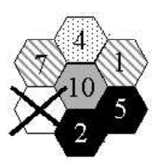


Figure 3. Case A3.

No transmitters in these cells can use the channels 8, 9, 10, 11, since the cells are within $2s$ of the center transmitter operating on 10; and none can be assigned channel 3, as we showed in Case A1. We are left with six usable channels, 1, 2, 4, 5, 6, 7, all of which we must use, since six distinct channels are required to fill the concentric. Channel 4 must be assigned to one of the cells, as in the dotted cell in **Figure 3**, and the striped cells neighboring it must contain channels 1 and 7. The $2s$ constraint requires that the cell with channel 5 can be adjacent only to the cells using channels 1 and 2, so the 5 and 2 must be added as shown in the figure (in black). However, we cannot assign channel 6 to the remaining cell, because that would violate the $2s$ requirement (since $7 - 6 = 1 < k$). It follows that no transmitter may be assigned channel 10.

We interrupt the flow of the argument to establish a claim that we need.

Claim: Any network of transmitters can be renumbered so that some transmitter operates on channel 1.

Proof: Suppose that there is a set of channels where no transmitter operates on channel 1. Let a be the smallest channel. As in Case 2, we renumber every channel m, this time as $m-a+1$. This new numbering preserves differences between channel assignments, so it still satisfies the difference constraints. Moreover, it contains channel 1 and all numbers in it are positive integers.

So we can assume that some transmitter is assigned channel 1. Consider the first concentric around this transmitter. No transmitters in these cells can use channels 1, 2, or 3, because the cells are within $2s$ of the transmitter operating on 1, and none can be assigned channels 9 or 10, as we showed in Cases A2 and A3. We are left with six usable channels, 4, 5, 6, 7, 8, 11, all of which we must use since six distinct channels are required to fill the first concentric. Channel 6 must be assigned to one of the cells, but there are not two numbers remaining in the list (4, 5, 7, 8, 11) that differ from 6 by more than $k = 3$. Therefore, it is impossible to complete the concentric in a way that satisfies the $2s$ constraint. This contradicts our supposition that we could assign channels using numbers between 1 and 11. Hence, when $k = 3$, the span must be greater than 11. □

Assertion B: *For $k = 3$, the span is 12.*

Proof B: The span is at least 12; we construct a solution that realizes 12. As was the case for $k = 1$ and $k = 2$, there exists a sequence of integers that, when applied in a series of adjacent, offset columns, produces a satisfactory network on an infinite plane (as in **Figure 2** for $k = 2$). The central column for $k = 3$ is the sequence 1, 8, 3, 10, 5, 12, 7, 2, 9, 4, 11, 6 repeated over and over. The adjacent column to one side is the same sequence shifted down 4, and the adjacent column to the other side is the sequence shifted up 4. As before, this pattern can be repeated indefinitely, and since the neighborhood of each number is exactly the same, the conditions are met. Therefore, for $k = 3$, the span is 12. □

$k > 3$

We prove that for $k > 3$, the span is $2k + 7$. We must first prove that no assignment with the channels 1 through $2k + 6$ satisfies the constraints. This proof requires a detailed analysis. [EDITOR'S NOTE: We omit the details of this analysis.] We must also show that there is a configuration of channels with span $2k + 7$ satisfying the constraints; **Figure 4** shows our solution. The same rhombus pattern is repeated over and over, tiling the plane (for example, the dotted, striped, and gray parallelograms are all identical copies). This rhombus consists of the numbers 1, 2, 3, $k+3$, $k+4$, $k+5$, $2k+5$, $2k+6$, and $2k+7$. As before, the neighborhood of each 1 is identical, and we can see that it satisfies the constraints, as do the neighborhoods of the other 8 cells. Therefore, $2k + 7$ is the span for all $k > 3$. $\square$

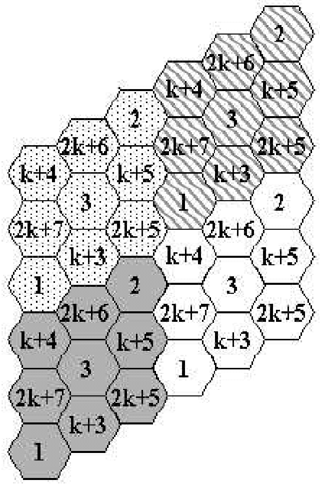


Figure 4. Configuration that tiles the plane to show that $2k + 7$ channels will work.

More Generalizations

We generalize the $4s$ constraint to: Transmitters $4s$ apart must have channels m apart, $m \leq k$. While we do not determine the span in this general case, we deduce bounds on it: It must lie between $1 + 2k + 4m$ and $1 + 2k + 6m$.

First, suppose that we can make correct assignments using only channels $1, \ldots, 2k+4m$. Let $A = \{1, 2, \ldots, k\}$ and $B = \{k+4m+1, k+4m+2, \ldots, 2k+4m\}$. All numbers in A are within k of each other, as are all numbers in B. Consider three cells that share a common vertex. At most one of these three can be assigned an element of A, and at most one can be assigned an element of B; so the third must be assigned a channel n between $k + 1$ and $k + 4m$. Consider the first concentric about this central cell with channel n. We need 7 numbers to make enough assignments to fill this first concentric (including the central cell, n). We label these in increasing order: $x_1 < \cdots < x_7$.

Case 1: n one of x_2, x_3, x_4, x_5, x_6. Since all of these transmitters are within $4s$ of each other, each of the gaps between x_1 and x_2, between x_2 and x_3, and so on must contain at least $m - 1$ numbers; two of these six gaps (the two around n), must contain at least $k - 1$ numbers. Summing up the seven channels in the first concentric and the channels in the gap, we need $7 + 4(m - 1) + 2(k - 1) = 1 + 2k + 4m$ channels, which contradicts our earlier assumption that we could make the assignments using only $2k + 4m$.

Case 2: $n = x_1$. This means that n is the smallest of the numbers. We still have one gap of size $k - 1$ (between n and x_2), and the rest are of size $m - 1$. Furthermore, since n is chosen so that it is at least $k + 1$, there are k channels below it. Therefore, we need $k + 7 + 1(k - 1) + 5(m - 1) = 2k + 5m > 2k + 4m$

channels, which contradicts our assumption.

Case 3: $n = x_7$. This is the same as $n = x_1$, except that n was chosen to be at most $k + 4m$. Therefore, we need at least a span of $1 + 2k + 4m$ channels to make correct assignments.

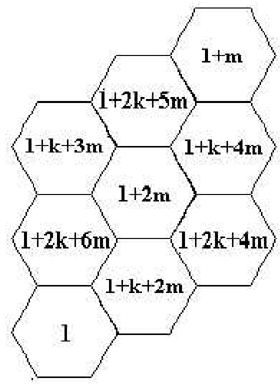


Figure 5. A rhombus that tiles the plane and uses only integers between 1 and $1 + 2k + 6m$.

A generalized network using only the integers between 1 and $1+2k+6m$ is shown in **Figure 5**. The situation is analogous to the $k > 3$ case discussed earlier. This time, we have a rhombus that tiles the plane, with channels assignments $1, 1+m, 1+2m, 1+k+2m, 1+k+3m, 1+k+4m, 1+2k+4m, 1+2k+5m$, and $1+2k+6m$. As before, we can check the neighborhood of each channel to make sure that it satisfies the constraints across all values of m and k.

To assess whether $1+2k+6m$ is a good upper bound for the span, consider how much smaller the span could be. There is a lower bound of $1 + 2k + 4m$, so our upper bound is not more than $2m$ greater than the span. Furthermore, for $m = 1$, we have $1 + 2k + 6m = 2k + 7$, which is exactly the span for $k > 3$.

Most important, the pattern we offer in this section provides a surprisingly efficient way to generate assignments for any sized grid, based on k and m: One need simply construct a rhombus of nine hexagons and tile the grid. In summary, though we have not proven that the span is $1 + 2k + 6m$, that expression appears to be a close approximation.

More Layers of Interference

We consider what happens if there are three levels of interference. We construct a method for deriving assignments that satisfy all the conditions:

- $2sk$ **constraint:** Channel assignments for transmitters within a distance of $2s$ of each other must differ by k.
- $4sm$ **constraint:** As above, but within a distance $4s$ they must differ by m.
- $6sn$ **constraint:** As above, but within a distance $6s$ they must differ by n.

We require that $n \leq m \leq k$.

We build up this assignment from a 2-level interference assignment. **Figure 6** shows a triangular lattice that results from drawing lines between centers of adjacent hexagons, while **Figure 7** shows a triangular lattice that connects only some of the cells. **Figure 7** looks identical to **Figure 6** but on a larger scale. In **Figure 7**, the dotted cells are within $4s$ of the central gray cell but more than $2s$ away, while the striped cells are within $6s$ of the central cell but more than $4s$ away.

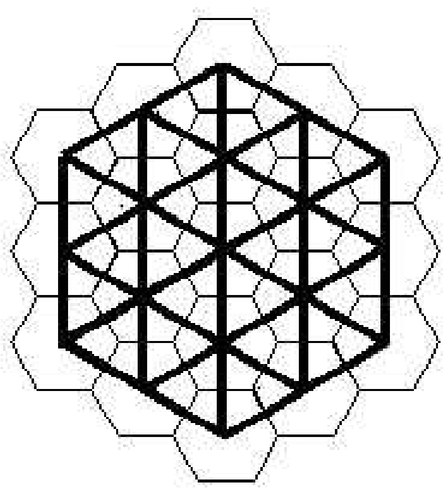


Figure 6. Triangular from drawing lines between centers of adjacent hexagons.

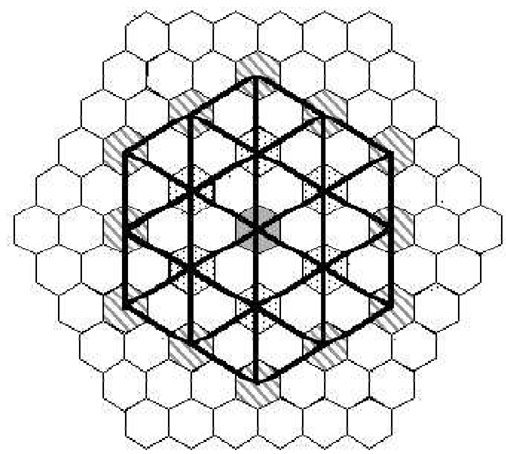


Figure 7. Triangular lattice that connects only some of the cells.

Suppose that an assignment satisfies both the $2sm$ and the $4sn$ constraints. If we assign these channels to the vertices of the lattice of **Figure 7**, they will meet the $4sm$ and $6sn$ constraints. **Figure 8** shows how we can overlap three lattices (light gray, dark gray, and black) such that every hexagon is centered on a vertex of one of the three lattices.

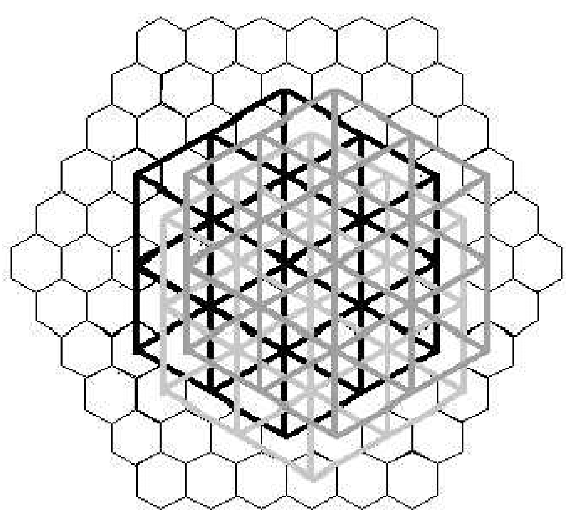


Figure 8. How the three lattices can be overlapped so that every hexagon is centered on a vertex of one of the three lattices.

We give the details. Suppose that the assignments on the light gray lattice use 1 through L and satisfy the $2s$ and $4s$ constraints. We label the cells on the dark gray lattice with $k+L$ to $k+2L-1$ (by simply adding $k+L-1$ to each channel, following the same assignment). We label the cells on the black lattice with $2k+2L-1$ to $2k+3L-2$ (by adding $2k+L-2$ to each channel).

Cells on different lattices have channels at least k apart. Cells on the same lattice are more than $2s$ apart; if they are less than $4s$ apart, their channels differ by m; and if they are less than $6ns$ apart, their channels differ by n. Thus, the assignment meets all the constraints, with maximum channel $2k+3L-2$.

Figure 9 gives a practical example of this method. Suppose that we seek a configuration for which channel assignments for transmitters within $2s$ of one another must differ by at least 3, those within $4s$ of one another must differ by at least 2, and those within $6s$ of one another must differ by at least 1. We use the configuration derived earlier (using the $2s2$ and $4s1$ constraints), which uses 9 integers. The gray vertices in **Figure 9** use the integers 1 through 9, the dotted vertices use 12 through 20, and the striped vertices use 23 through 31.

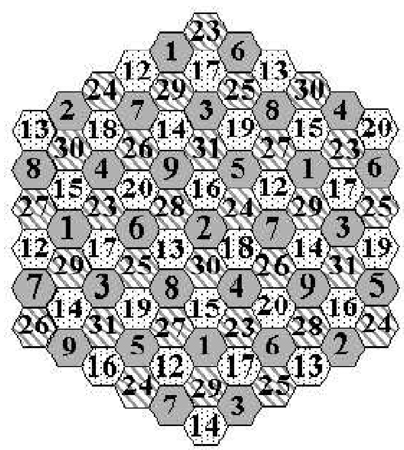


Figure 9. Example of labeling constructed to satisfy prescribed constraints.

We can apply this process to get an assignment configuration that satisfies the $2sk$, $4sm$, and $6sn$ constraints, for arbitrary k, m, and n. We have already shown that there exists an assignment configuration satisfying the $2sm$ and $4sn$ requirements whose maximum integer is $1+2m+6n$. Using the above method, we obtain a configuration that satisfies the $2sk$, $4sm$, and $6sm$ requirements, and its largest integer (substituting $1+2m+6n$ for L) is $2k+3(1+2m+6n)-2 = 1+2k+6m+18n$.

Does this method produces efficient configurations? That is, is the maximum integer that it obtains close to the actual span? While we have no proof, we suggest why it is an efficient method. We use the method to move from the 2-layer interference to the 3-layer interference, but we could have used it to move from 1 layer to 2 layers. So let's use this method to generate an assignment configuration with $2sk$ and $4s1$ constraints. We begin by finding the span when the only constraint is the $2s1$ constraint (i.e., that adjacent cells must have different channels). This is clearly accomplished by the sequence 1, 2, 3 repeated in a central column, shifted down two in the adjacent column to the right, and shifted up two in the adjacent column to the left.

If we use our method to construct a configuration with $2sk$ and $4s1$ constraints, its maximum channel assignment (substituting 3 for L) would be $2k+3\times 3-2 = 2k+7$. This is the span for $k > 3$. Therefore, our method generates a 2-layer interference from a 1-layer interference efficiently. It is reasonable, therefore, to expect that it also generates 3 layers from 2 fairly efficiently.

It is possible to expand to even higher layers of interference using our method. For example, in **Figure 10**, the striped dots are all between $9s$ and $6s$ of the gray cell. A 3-layer interference assignment configuration on the lattice gray, dotted, and striped cells can produce an assignment configuration on the whole grid, with constraints for $2s$, $4s$, $6s$, and $9s$.

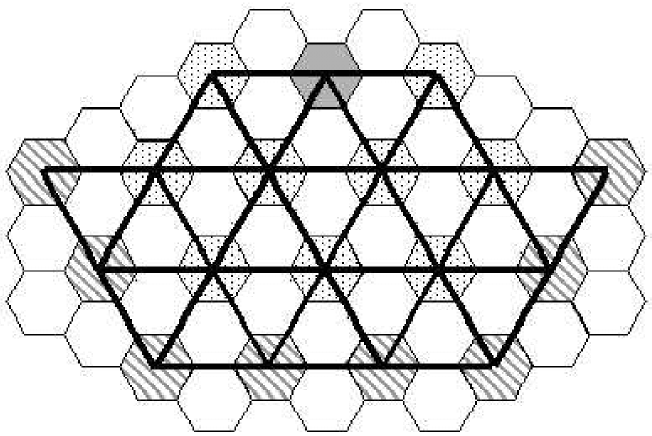


Figure 10. Example of labeling satisfying higher layers of interference.

Students Clamor for Bandwidth Optimization

WINSTON-SALEM, Feb. 30 — A team of three college students testified today before the Congressional Subcommittee on Bandwidth Regulation, revealed key research findings that may unleash a flood of proposed legislation designed to boost the efficiency of the information economy.

Several months ago, the Subcommittee issued a challenge to the world's mathematicians to find a method by which the United States can conserve its bandwidth efficiently. Yesterday, the three students, all from Wake Forest University here in Winston-Salem, stunned the world with their solution to how to assign radio channel frequencies. They found patterns of frequency assignments that maximize efficiency while maintaining the quality of the signals. The new discovery may be a breakthrough for frequency assignments for TV, radio, wireless modems, and cellular phones.

Furthermore, the students devised methods that could help the government determine the optimum number of channels for a given area, based on the likelihood of interference between channels. These models may have far-reaching implications, possibly affecting how the Federal Communication Commission assigns radio channels in the future.

Since its inception, the Congressional Subcommittee on Bandwidth Regulation has sought to make sure that all available parts of the spectrum are conserved for government, commercial, and private use. "We minimize interference by careful regulation of station licensing," commented Subcommittee Chairwoman Jane Doe (D–NY). "And this new discovery will help. Simply put, if we don't waste what we have, there will be more left over to sell, which could mean lower taxes."

In spite of the potential benefits of the discovery, there was some dissension about implementing policies based on it. "While I agree that the patterns that these kids have generated are quite beautiful from a mathematical standpoint," commented Sen. Laissez Fair (R–TX), "government regulation is not necessary in an industry that tends to regulate itself. After all, radio stations tend to space themselves out naturally."

Others at the hearing disagreed with Mr. Fair, noting that all popular stations seemed to have converged inexplicably to the high side of the FM band, leaving large portions of the FM spectrum unused ("except for that useless government-welfare National Public Radio down at the bottom of the band somewhere," retorted Mr. Fair). Ms. Doe responded that "The economic consequences are important. We are certainly going to recommend legislation to optimize channel assignments on portions of the bandwidth that are already in use. And future assignments should follow the patterns that these bright young mathematicians have discovered."

However, industry sources signaled that they are vehemently opposed to radio and TV stations (including satellite TV) being be forced to change frequencies to comply with any new efficiency standards. To this industry reaction, one of the student researchers commented, "Our goal isn't to force our model of efficiency on a market that has been functioning for decades. We're just trying to help the government plan for future expansion."

The Subcommittee on Bandwidth Regulation was initially formed in response to the public's concerns surrounding the notorious HDTV "bandwidth heist," which became a popular issue for the Bob Dole campaign in the 1996 presidential election. With widespread support from his party, Dole promised to auction off the new HDTV broadcast spectrum rather than give it away to TV networks interested in converting to HDTV. Sen. John McCain (R–AZ), a front-runner in this year's Republican presidential primary and Chairman of the Senate Commerce, Science and Transportation Committee, has estimated that such an auction would bring in over $70 billion that could help to reduce taxes. "After all, the public airwaves are owned by the American people and managed by our government," said a McCain spokesperson.

Radio Channel Assignments

Justin Goodwin
Dan Johnston
Adam Marcus
Washington University
St. Louis, MO

Advisor: Hiro Mukai

Summary

We use mainly combinatorial methods to estimate and prove bounds for various cases, concentrating on two levels of interference. We use the concept of a *span*, the minimum largest channel among assignments that satisfy the constraints.

For Requirements A and B, the span is 9. For Requirement C, the span is 7 when $k = 1$, 9 when $k = 2$, 12 when $k = 3$, and $2k + 7$ for $k \geq 4$.

For Requirement D, we present the results in a table (**Table 2**). Some of our results improve on upper bounds in Shepherd [1998].

Only regular transmitter placement needs to be considered; irregular placement can be accommodated by making the hexagons so small that the transmitters are in regular placement, with the bounds adapted correspondingly.

For Requirement E, we discuss both the limitations of our model and its ability to produce an upper bound for any situation.

Definitions

- Let s denote the length of a side of a hexagon. Then the distance from the center of one hexagon to the center of an adjacent hexagon is $s\sqrt{3}$.
- A *region* is a collection of hexagons, finite or otherwise.
- For u and v hexagons in a region $\mathcal{X}$, let $D(u, v)$ be the minimum number of hexagons (including the first but not including the last) that one must pass through to move from u to v in region $\mathcal{X}$. Set $D(u, u) = 0$. So, for example, the stipulation in the problem that any two different transmitters

The UMAP Journal 21 (3) (2000) 369–378.

(in hexagons u and v) are within distance $2s$ is equivalent to $D(u, v) \leq 1$. Similarly, two hexagons are within distance $4s$ if and only if $D(u, v) \leq 2$.

- Let T be the the portion of a plane that includes a hexagon u along with all hexagons v such that $D(u, v) \leq 3$ (**Figure 1**).

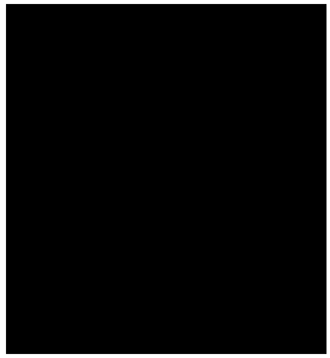


Figure 1. Region T.

- Let R be any planar hexagonal grid that contains T.

- Let k_i be the minimum allowed difference in channels of two hexagons u and v in a region R that have $D(u, v) = i$. For example, if $k_1 = 2$ and $k_2 = 1$, then transmitters in hexagons u and v that are adjacent must have channels that differ by at least 2. If the transmitters in hexagons u and v are two hexagons apart (i.e., $D(u, v) = 2$), then their channels must not be the same.

- Let C be a function from the hexagons in a region R to the positive integers. Given a set of constraints, call C a *channel assignment* to R under those constraints if C maps the hexagons to an allowed set of frequencies.

- The *width of the interval* of the frequency spectrum in region R is the largest channel used. The minimum width over all channel assignments of a region R is the *span*.

- Let the function $S(l_1, l_2, \ldots, l_n)$ of a region R be the span under the restrictions that $k_i = l_i$ for all i from 1 to n.

- For a given $k_1 \geq 4$, define the set

$$\mathcal{N}_k = \{1, 2, 3, k+3, k+4, k+5, 2k+5, 2k+6, 2k+7\}$$

as the *channel assignment set*. That is, for a region R, $C(R) \subseteq \mathcal{N}_k$.

Solution

We are concerned with planar regions that expand out in every direction infinitely or else are finite. First, we prove some general results.

Lemma 1. *Let M be any positive integer. If $S(k_1, k_2, \ldots, k_n) = L$, then*

$$S(k_1, k_2, \ldots, k_{i-1}, k_i + M, k_{i+1}, \ldots k_n) \leq L + M \left(\left\lceil \frac{L}{k_i} \right\rceil - 1 \right).$$

Proof: Let C_1 be an assignment of channels on the region R with span L and satisfying the given constraints. We construct an assignment that satisfies the new constraints, with the desired largest channel. Define a new channel arrangement C_2 as follows:

$$C_2(u) = C_1(u) + M \left(\left\lceil \frac{C_1(u)}{k_i} \right\rceil - 1 \right).$$

To see that the new constraints are satisfied, notice that

$$|C_2(u) - C_2(v)| \geq |C_1(u) - C_1(v)|;$$

so all the constraints for k_j, $j \neq i$ are still satisfied. Furthermore, if

$$|C_1(u) - C_1(v)| \geq k_i,$$

then

$$|C_2(u) - C_2(v)| = |C_1(u) - C_1(v)| + M.$$

This is because if $|C_1(u) - C_1(v)| \geq k_i$, then

$$\left\lceil \frac{C_1(u)}{k_i} \right\rceil \neq \left\lceil \frac{C_1(v)}{k_i} \right\rceil.$$

This demonstrates that the constraint for the new value of k_i is now satisfied. Thus, the only channels used are of the form

$$C_1(u) + M \left(\left\lceil \frac{C_1(u)}{k_i} \right\rceil - 1 \right) \leq L + M \left(\left\lceil \frac{L}{k_i} \right\rceil - 1 \right).$$

Therefore, the channel assignment that we have constructed is valid, and we have further shown the desired feature that

$$S(k_1, k_2 \ldots k_{i-1}, k_i + M, k_{i+1}, \ldots k_n) \leq L + M \left(\left\lceil \frac{L}{k_i} \right\rceil - 1 \right). \qquad \square$$

Lemma 2. *On any region R containing T, $S(4,1) > 14$.*

[EDITOR'S NOTE: We omit the proof.]

Lemma 3. *On any region R containing T, $S(3,1) > 11$ and $S(2,1) > 8$.*

Proof: If $S(3,1) = L \leq 11$, then by **Lemma 1** we know that

$$S(4,1) \leq L + \left\lceil \frac{L}{3} \right\rceil - 1 \leq 11 + \left\lceil \frac{11}{3} \right\rceil - 1 = 14,$$

which is a contradiction to **Lemma 2**. Similarly, if $S(2,1) = L \leq 8$, then by **Lemma 1** we have

$$S(3,1) \leq L + \left\lceil \frac{L}{2} \right\rceil - 1 \leq 8 + \left\lceil \frac{8}{2} \right\rceil - 1 = 11,$$

which violates what we just proved. $\square$

Lemma 4. *If $l > 4$, then for region T, $S(l,1) > 2l + 6$.*

[EDITOR'S NOTE: We omit the proof.]

The proof works for any region, finite or infinite.

$k_2 = 1$

For any two hexagons u and v, if $D(u,v) = 1$, then their channels differ by at least k ($k_1 = k$) for any positive integer k, and $k_2 = 1$. With this generalization, we would like to see how the span relates to k_1.

Lemma 5. *For $k_1 \geq 4$, a width of the interval of the frequency spectrum in region R is less than or equal to $2k_1 + 7$.*

Proof: By induction. We use the set defined in **Definition 9**. First we show that $k_1 = 4$ satisfies **Lemma 5**. If $k_1 = 4$, then for all u, v such that $D(u,v) \leq 1$ we have $|C(u) - C(v)| \geq 4$, by definition. By **Lemma 2**, we have $S(4,1) > 14$. To see that for $k_1 = 4$ a frequency width is $2(4)+7 = 15$, use the channel assignment set $\mathcal{N}_4 = \{1,2,3,7,8,9,13,14,15\}$. As shown in **Figure 2**, the channel assignment set satisfies the constraints.

Also, the channel assignments tessellate and the resulting pattern always meets the constraints. To see this, translate the channel assignments from A to B. After translation, we have a repeated pattern with no gaps and the constraints still hold. Now, instead translate from A to C, and again we have a repeated pattern with no gaps while keeping all constraints. Since these are the only two possible kinds of translation, we have shown that the pattern is a tessellation. Since the maximum channel assigned is 15, the width of the frequency spectrum is $15 = 2(4) + 7$.

Next, let k be any integer such that $k \geq 4$, assume that **Lemma 5** holds true for $k_1 = k$ with channel assignment set $\mathcal{N}_k$. This generates a tessellation as illustrated in **Figure 3**. It is easy to see that this pattern tessellates and meets the constraints.

Now we need to prove that **Lemma 5** holds for $k_1 = k + 1$. To do so, we generate a tessellation pattern from $\mathcal{N}_{k+1}$ in region R that satisfies the constraints. In our hypothesis that **Lemma 5** holds for k, we replace all k with $k+1$ in **Figure 3**. The result is the tessellation in **Figure 4**, which meets all the constraints. The maximum frequency used is $2k + 9 = 2(k+1) + 7$; that is, for $k+1$ the width of the interval of the frequency spectrum is $2(k+1) + 7$. Since this matches our inductive hypothesis for $k+1$, we have proven the lemma. $\square$

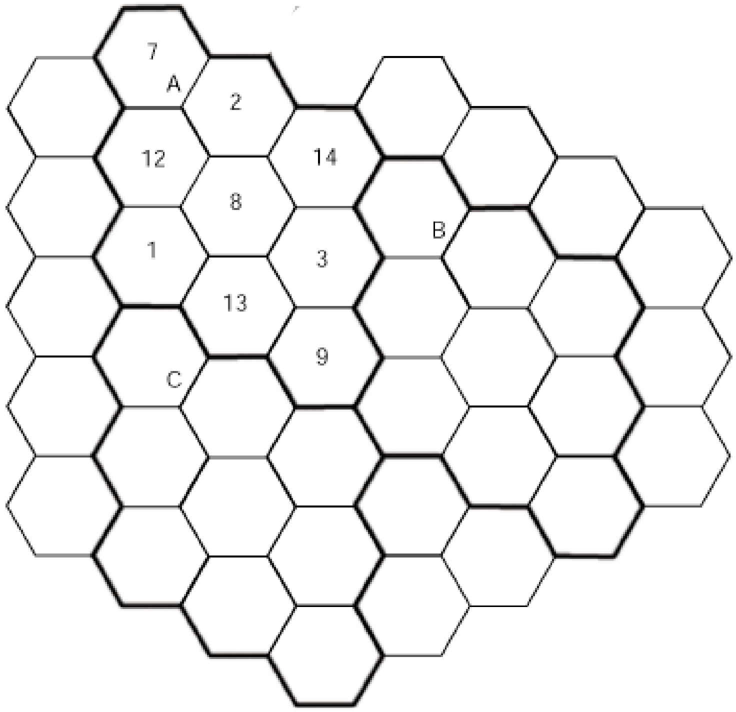


Figure 2. Channel assignment for $k_1 = 4$.

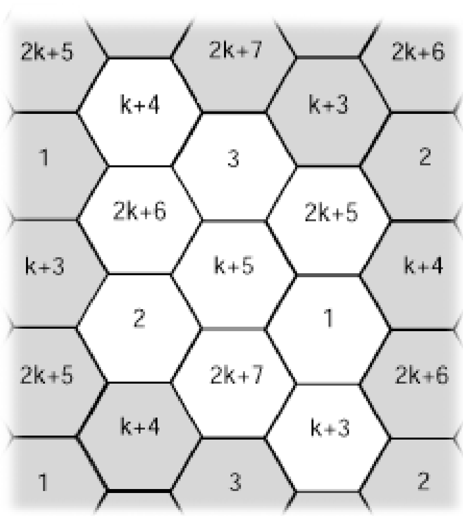


Figure 3. Channel assignment for $k_1 = k$.

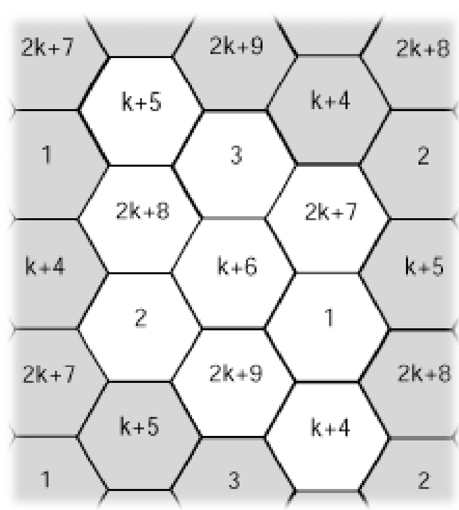


Figure 4. Channel assignment for $k_1 = k + 1$.

Lemma 5 is a very nice result. It gives a way of constructing a tessellation under the constraints that $k_1 \geq 4$ and $k_2 = 1$, and we can make this pattern using the channel assignment set $\mathcal{N}_{k_1}$. Most important, we can assign the channels with a frequency width of $2k_1 + 7$. Next, we prove that this width is a lower bound for any $k_1 \geq 4$.

Theorem 1. *Let $k_1 \geq 4$. Then $S(k_1, 1)$ of a region R is $2k_1 + 7$.*

[EDITOR'S NOTE: We omit the proof.]

With **Theorem 1** and **Lemma 5**, we know how to form a repeating pattern for the given constraints, and we also know the span over the region R. A very nice outcome from these results is that for *any* $k_1 \geq 4$, we can choose nine connected hexagons and produce a channel assignment with $S(k_1, 1) = 2k_1 + 7$. By looking at **Figure 3**, we can see that for large k values we have a larger spread in frequencies; that is, for larger k_1 we have a more efficient system of transmitters in terms of interference because the frequency width is large.

We now attend to $k_1 = 2$ and $k_1 = 3$.

Theorem 2. *For $k_1 = 2$, the channel assignment set is*

$$\mathcal{N}_2 = \{1, 2, 3, 4, 5, 6, 7, 8, 9\},$$

with $S(2, 1) = 9$.

For $k_1 = 3$, the channel assignment set is

$$N_3 = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\},$$

with $S(3, 1) = 12$.

Proof: By **Lemma 3**, we know that $S(2, 1) > 8$ and $S(3, 1) > 11$. In **Figure 5**, we have a tessellation pattern for $k_1 = 2$ with channel assignment set as C_2. By inspection, $S(2, 1) = 9$, the lowest possible value. For $k_1 = 3$, we have a similar argument, only we use the channel assignment set C_3. By inspection of **Figure 6**, $S(3, 1) = 12$, the lowest possible value. $\square$

$k_1 = k$ and $k_2 = 0$

The values in **Figure 7** meet the constraints. Therefore, the span over a region R for this case is $2k + 1$. To see this, try for $k - 1$; then the channel assignment set is $\{1, k, 2k - 1\}$, but k and 1 must be at least k apart. Hence, $2k + 1$ is the span.

$k_1 = k_2 = k$

The values in **Figure 8** meet the constraints. Hence, the span over a region R is $6k + 1$. To see this, as above try for $k - 1$; then we have a contradiction in **Figure 7** with the hexagons containing 1 and $(k - 1) + 1 = k$. Therefore, $2k + 1$ is the span.

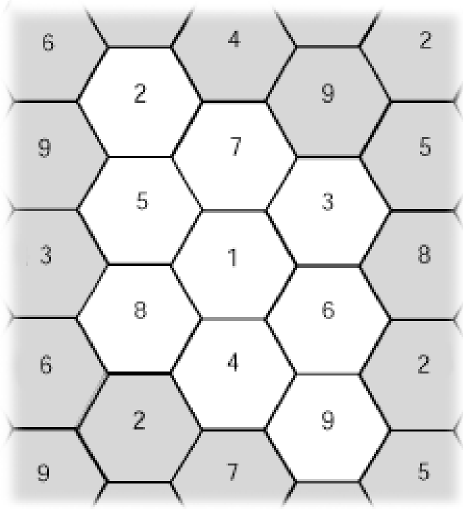


Figure 5. Channel assignment for $k_1 = 2$.

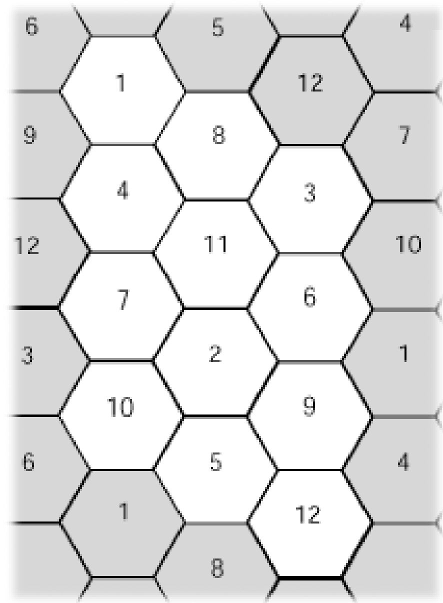


Figure 6. Channel assignment for $k_1 = 3$.

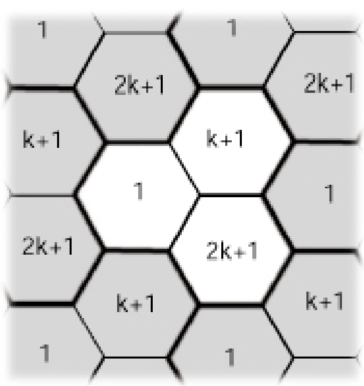


Figure 7. Channel assignment for $k_1 = k$ and $k_2 = 0$.

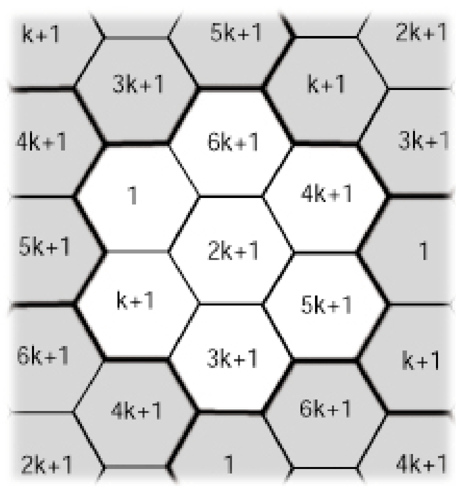


Figure 8. Channel assignment for $k_1 = k_2 = k$.

General Case

In this section, k_1 and k_2 can be any positive integers.

Theorem 3. *Let R be a region that contains region T and let $k_1 \geq 4k_2$. Then*

i) *If k_2 divides k_1, then $S(k_1, k_2) = 2k_1 + 6k_2 + 1$.*

ii) *If $k_1 > 6k_2 + 1$, then $S(k_1, k_2) = 2k_1 + 6k_2 + 1$.*

[EDITOR'S NOTE: We omit the proofs.]

Theorem 4. *Let $3k_2 \leq k_1 \leq 4k_2$. Then $S(k_1, k_2) \leq 3k_1 + 2k_2 + 1$.*

Proof: By construction. Consider the tiling in **Figure 9**. As long as $3k_2 \leq k_1 \leq 4k_2$, the channel assignment holds. Then by construction,

$$S(k_1, k_2) \leq 3k_1 + 2k_2 + 1.$$

As shown by the highlighted tiles, this tiling works only if $2k_2 + 1$ and $k_1 + 1$ differ by at least k_2 (by definition of k_2).

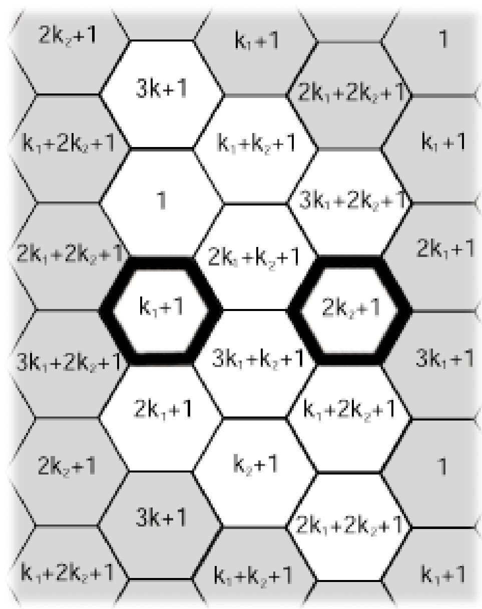


Figure 9. Channel assignment for **Theorem 4**.

It follows that

$$\begin{aligned} (k_1 + 1) - (2k_2 + 1) &\geq k_2, \\ k_1 - 2k_2 &\geq k_2, \\ k_1 &\geq 3k_2. \end{aligned}$$

Yet we know from **Theorem 3** that for $k_1 \geq 4k_2$ we have a strict lower bound; therefore, we must have a strict upper bound, that is

$$k_1 \leq 4k_2.$$

Hence, we have that if $3k_2 \leq k_1 \leq 4k_2$, then $S(k_1, k_2) \leq 3k_1 + 2k_2 + 1$. □

Conclusion

We summarize specific proved results in **Table 1**. For the cases $k_2 = 2$ and $k_1 = 9, 11, 13$, we are unable to determine $S(k_1, k_2)$. However, we find bounds for those values by **Lemma 2**.

Table 1.

Compilation of spans for different values of k_1 and k_2.

k_1	k_2	$S(k_1, k_2)$
1	1	7
2	1	9
3	1	12
4	1	15
5	1	17
$l > 5$	1	$2l + 7$
2	2	13
3	2	17
4	2	17
5	2	21
6	2	23
7	2	26
8	2	29
9	2	30, 31 or 32
10	2	33
11	2	34, 35 or 36
12	2	37
13	2	39 or 40
$l > 13$	2	$2l + 13$

Table 2.

General results for values of k_1 and k_2.

Constraints	Span
any k_1 , $k_2 = 0$	$2k_1 + 1$
$k_1 = k_2$	$6k_1 + 1$
$k_1 = 2$, $k_2 = 1$	9
$k_1 = 3$, $k_2 = 1$	12
$k_1 \geq 4$, $k_2 = 1$	$2k_1 + 7$
$k_1 \geq 4k_2$	$\leq 2k_1 + 6k_2 + 1$
$3k_2 \leq k_1 \leq 4k_2$	$\leq 3k_1 + 2k_2 + 1$
$k_1 > 4, k_2 = 1$	$> 2k_1 + 6$
$3k_2 \geq 2k_1$	$4k_1 + 3k_2$

Table 2 has our general results. The last row was proven not by us but by Mark Shepherd [1998]. For selected values of k_1 and k_2, we establish the span of an arbitrary planar hexagonal region that includes T. For all combinations, we can find a pattern that repeats—that is, we can find a tessellation of frequencies. This is a major result, because we know how to construct a frequency assignment based on the values of k_1 and k_2 through a simple formula, as shown in **Figure 4** for $k_1 \geq 4$ and $k_2 = 1$.

News Release: Mathematicians Help Clear Out Airwaves

Last week mathematical researchers at Washington University in St. Louis announced that they have improved the current method of assigning radio frequencies, such as the channel of your favorite station. With the increase in wireless communication, it has become more important than ever to assign frequencies efficiently while avoiding interference as much as possible.

The mathematicians made use of a certain type of pattern, made popular by the contemporary artist M.C. Escher, called a tessellation. These patterns are carefully constructed to cover an entire region without leaving any gaps. The new results show how to assign channel frequencies to regions in a tessellation so as to minimize several kinds of interference from nearby stations. The work extends efforts currently in progress at Oxford University in England.

"Our work is quite general," commented one of the researchers. "It applies regardless of geographical situations, such as differences in altitude or other natural phenomena."

Radio listeners have nothing to fear from these new developments; the frequency of your favorite station is unlikely to change. The new results will help long-term planning by engineers, operators of cell-phone services, and government regulators. "In the future you won't have the kind of interference that causes someone to flip to Rush Limbaugh's channel but end up instead with Howard Stern. We're just making sure that listeners get to hear Rush say his whole two cents worth."

Reference

Shepherd, Mark. 1998. Radio channel assignment. Ph.D. thesis, Merton College, Oxford University. `http://www.maths.ox.ac.uk/combinatorics/thesis.html`.

Author/Judge's Commentary: The Outstanding Channel Assignment Papers

Jerrold R. Griggs
Dept. of Mathematics
University of South Carolina
Columbia, SC 29208
griggs@math.sc.edu
homepage: http://www.math.sc.edu/~griggs/

Background

This 2000 MCM Problem, which I wrote, is based on a subject of considerable current interest to mathematicians and communications engineers. The original "channel assignment problem" has a long history. The problem is to assign an integer channel to each transmitter in a network, with the condition that the absolute difference between channels for two nearby transmitters must not belong to a certain set T that arises from interference considerations (see Hale [1980] for motivation). A feasible assignment can be obtained with channels far apart, but this is highly inefficient. Typically, a frequency band that spans the assigned channels is allocated to the network; the wider the band, the more it costs. The problem, then, is to minimize the "span" of the assignment, which is the difference between the maximum channel and the minimum channel.

This problem is modeled nicely with graph theory by letting each transmitter correspond to a vertex, with edges corresponding to pairs of nearby transmitters. The problem becomes a special vertex-coloring problem, owing to the set T of forbidden differences [Cozzens and Roberts 1982]. Among the methods that come into play are number theory (in the case of complete graphs [Griggs and Liu 1994]) and the complexity of graph homomorphisms.

The UMAP Journal 21 (3) (2000) 379–386.

History of the Problem

In 1987, I heard from Fred Roberts [1987] about a variation of the T-coloring problem in which channels for two nearby transmitters, say within distance s, must differ by at least two, while those within distance $2s$ must differ by at least one. The network may have thousands of transmitters. Again we seek to minimize the span of a feasible channel assignment.

In considering this problem, I realized that it no longer translates directly into a graph theory problem by assigning vertices to transmitters and edges to pairs at distance at most s: Although two transmitters at distance at most s correspond to adjacent vertices, a pair of transmitters at distance between s and $2s$ corresponds to a pair of vertices at distance two in the graph only when some third transmitter is within distance s of both of them. (Note that the *distance* between two vertices in a graph is the number of edges in a shortest path between them.) In fact, the vertices for two transmitters at distance between s and $2s$ may not even be connected in the graph.

Nonetheless, it is clear that for the real problem it is useful to understand the natural graph analogue, which is to find the minimum span for the integer labelings of a graph such that labels for vertices at distance one (resp., two) differ by at least two (resp., one). For the transmitter networks in Parts A and B of the contest problem this year, the associated graph problem is precisely of this type with one change: The span in the contest problem is one more than in the labeling problems in the literature.

Griggs and Yeh [1996] introduce this graph labeling problem and pose some fundamental questions about it. Included is the natural generalization of the graph problem in which there are multiple levels of spectral spreading interference: Given integers $d_1, \ldots, d_r$ we seek minimum span labelings such that for all i, the labels for any pair of vertices at distance i differ by at least d_i. Such a labeling is called an $L(d_1, \ldots, d_r)$*-labeling*.

Applications

While such problems are mathematically interesting, they have taken on greater importance in recent years due to their potential applicability to the design of mobile radio networks. Large areas are often covered by a network of regularly spaced transmitters such that the associated graph labeling problem exactly models the network problem. The most common design places the transmitters in a triangular lattice, so that the whole region can be tessellated by a honeycomb of hexagons, with each transmitter in the center of a hexagonal region that it covers. An early reference considering such a model is Gamst [1982]; and evidently MacDonald [1979], cited by contest teams, also does this. A group led by Robert Leese at Oxford has been prominent in this program in recent years [Leese 1999].

The Outstanding Papers

Requirement A

Part A of this year's contest problem is a basic instance of this application, a sizable array of transmitters with $d_1 = 2, d_2 = 1$. Feasible solutions are trivial to find, but working down to an optimal one requires some cleverness. We had expected many teams would enjoy working on this and that most would achieve the minimum span; this was indeed the case.

Requirement B

Part B extends the network of Part A to the whole plane. While one can solve Part A by trial-and-error or by a computer search (to obtain an optimal assignment and rule out smaller ones), Part B requires a method to keep going forever labeling this infinite network. Successful teams for Part B usually found a pattern (a strip or a tile of numbers) that could be repeated indefinitely and achieve the same span as the bounded array in Part A. One way is to label a strip by an appropriate ordering, say

$$1\ 3\ 5\ 7\ 9\ 2\ 4\ 6\ 8\ 1\ 3\ \ldots,$$

and then use the same strip shifted appropriately for the next row, and so on. Another perspective is to construct an appropriate tile of nine hexagons and replicate it. Judges were pleased to see papers, such as that of the team from California Polytechnic State University, that test a variety of heuristics to assign channels in Parts A and B, since such methods are needed for more general arrays and distance parameters. At least one paper, by the team from Wake Forest University, makes the interesting observation (with proof) that the optimal labeling for Parts A and B is essentially unique!

Requirement C

What is most remarkable is that several teams were able to solve Part C, in which the channel spread parameters for Parts A and B are extended to $d_1 = k, d_2 = 1$. One can give decent labelings for the bounded array in Part A and for the full plane in Part B that are not far off from the lower bounds that one can quickly derive. However, we had not completely solved the problem for general k before the contest. It seems to be a new result.

Requirement D

Part D is the open-ended generalization of the problem to general array configurations and multiple levels of interference. It has the most room for creativity and for model design. This part was expected to be the main point of differentiation among the entries. Judges were disappointed that most entries did not do much here—perhaps they ran out of time working on Parts A–C, where the assumptions and model are explicit. Weaker papers only considered, say, what happens if one transmitter is not at the center of its hexagon. But stronger papers gave this part considerable thought. Some considered general conditions with two levels of interference, such as the impressive results contained in the paper by the team from Washington University that nearly solve it. (In part, they built on the thesis of Mark Shepherd [1998].) The paper from the Wake Forest University team analyzes an assignment method for multiple levels of interference. Judges wanted to see teams use the real problem of wireless communication as motivation, such the choice of multiple-level distance parameters analyzed by the team from Lewis & Clark University. Some considered how to adapt their hexagonal lattice approach to other configurations of transmitters.

Requirement E

For Part E, judges wanted to see an article that conveys to the public the sense of the problem and the team's ideas on how to attack it. A particularly amusing article was crafted by a team from Harvey Mudd College whose entry received Honorable Mention.

General Remarks

Several teams located related results in the literature or the Web, particularly for the problem of *cyclic* labelings, where integers $\{1, 2, \ldots, n\}$ are used but the distance between two labels is measured modulo n, that is, by the shortest path on the circle labelled 1 through n. This approach can be used when a large number of channels must be assigned to each location: When a vertex receives label i, it is given all channels congruent to i modulo n. For two levels of interference ($L(d_1, d_2)$), this cyclic problem is solved in van den Heuvel et al. [1998]. However, this does not immediately solve the contest problem. A solution for general $L(d_1, d_2)$ of the (linear) contest problem remains to be found.

Teams typically found good labelings for the bounded array in Part A by trial and error or by exhaustive computer search, for small values of k, and identified patterns or tiles that could be extended to general k to yield good labelings for the bounded and unbounded arrays.

One cannot be certain that a labeling is optimal without proving that there is no labeling of smaller span. Also, it is by no means clear that there exist optimal labelings using a repeating pattern, although many teams seemed to assume this. Thus, it is not sufficient to check just labelings from a repeating pattern. (In fact, it would be very interesting if one could show that for all sets of distance parameters d_i that there is an optimal labeling of the plane built from a repeating pattern. This seems to be an open question.)

Judges favored papers that provide a *clear proof* that their labelings are optimal for general k. The best proofs that we read were impressive, such as the one by the team from the National University of Defence Technology. That paper is among those that made the interesting observation that for general k there is an optimal labeling for the arrays in Parts A and B that uses only nine different channels—which could be useful in some applications!

Related Research

Chang and Kuo [1992] made noteworthy progress on the original graph labeling problems for $d_1 = 2, d_2 = 1$ posed by Griggs and Yeh. Griggs and Yeh conjectured that every graph of maximum degree $\Delta \geq 2$ has an $L(2,1)$-labeling of span at most Δ^2; this bound is achieved by cycles. Their conjecture remains open, even for $\Delta = 3$. For the famous Petersen graph, in which every vertex has degree 3, the minimum span is 9, the conjectured maximum.

Georges and Mauro [1995] showed how the $L(2,1)$-labeling problem for general graphs G is equivalent to a path covering problem for the complement of G. Such problems are known to be difficult (consider the problem of whether a graph has a Hamilton path, for instance), and, indeed, it has been recently shown [Fiala et al., to appear] that determining whether a graph has a labeling with span at most k is NP-complete for all $k \geq 4$. A good general upper bound on the span in the case $L(p,q)$ has been given recently [van den Heuvel and McGuinness 1999] for general planar graphs of maximum degree Δ, by applying the methods of the proof of the Four Color Theorem.

Leese [1997] considers channel assignments for the hexagonal array of our problem that are obtained by tilings (periodic labelings). A wide range of applied references is provided in this paper. McDiarmid and Reed [1997] and Fitzpatrick et al. [2000] discuss algorithms for channel assignments for the hexagonal array of our problem in which each location must v must receive a specified number w_v of channels. Many papers seem to be emerging that employ familiar methods of discrete optimization to produce channel assignments of low span (not necessarily optimal). Contest teams discovered work that we were not aware of, by Hurley [n.d.] and by Smith and Hurley [1997], that uses heuristics and search methods, including tabu search and genetic programming; Hurley developed software for these problems. A new project by Leese [2000] tests a linear programming method based on column generation.

Conclusion

Returning to the contest problem, judges had hoped to see more entries employ such methods of discrete optimization on Part B. We also hoped that more effort would be spent considering the open-ended modeling Part D of the problem. It may simply be that teams found more direct analysis to be successful on the specific problem instances in Parts A, B, and C, and most of their energy was spent on tackling these parts. The Outstanding papers published here are among the very few that accomplished much with the extension to Part D.

Judges raised fewer concerns than in past years about specific missing elements in entries; but again this is likely because the model for Parts A, B, and C, which teams focused on, is clear-cut. In general, what judges particularly sought in winning papers was clarity, both in explaining their approach and in proofs, especially for the lower bound in Part C. Since space permits, I reproduce below the discussion in my Judge's Commentary last year [Griggs 1999] of crucial elements in an outstanding contest entry.

Crucial Elements in an Outstanding Entry

Here are some general tips that the judges feel apply to any contest problem.

- *Teams should attempt to address all major issues in the problem.* Projects missing several elements are eliminated quickly.
- *A thorough, informative summary is essential.* Papers that are strong otherwise are often eliminated in early judging rounds due to weak summaries. Don't merely restate the problem in the summary, but indicate how it is being modeled and what was learned from the model. The summary should not be overly technical.
- *Develop a model that people can use!* The model should be easy to follow. While an occasional "snow job" makes it through the judges, we generally abhor a morass of variables and equations that can't be fathomed. Well-chosen examples enhance the readability of a paper. It is best to work the reader through any algorithm that is presented; too often papers include only computer code or pseudocode for an algorithm without sufficient explanation of why and how it works.
- *Supporting information is important.* Figures, tables, and illustrations are very helpful in selling your model. A complete list of references is essential—document where your ideas come from.

References

Chang, G., and D. Kuo. 1992. Labelling graphs with a condition at distance 2. *SIAM Journal of Discrete Mathematics* 5: 586–595.

Cozzens, M.B., and F.S. Roberts. 1982. T-colorings of graphs and the channel assignment problem. *Congressus Numerantium* 35: 191–208.

Fiala, J., T. Kloks, and J. Kratochvíl. To appear. Fixed-parameter complexity of λ-labellings. *Discrete Applied Mathematics*.

Fitzpatrick, S., J. Janssen, and R. Nowakowski. 2000. Distributive online channel assignment for hexagonal cellular networks with constraints. Preprint, January 2000.

Gamst, A. 1982. Homogeneous distribution of frequencies in a regular hexagonal cell system. *IEEE Transactions on Vehicular Technology* VT-31: 132–144.

Georges, J.P., and D.W. Mauro. 1995. Generalized vertex labelings with a condition at distance two. *Congressus Numerantium* 109: 141–159.

Griggs, Jerrold R. 1999. Judge's commentary: The outstanding lawful capacity papers. *The UMAP Journal* 20 (3): 331–333.

________, and D. D.-F. Liu. 1994. The channel assignment problem for mutually adjacent sites. *Journal of Combinatorial Theory A* 68: 169–183.

________, and R.K. Yeh. 1996. The $L(2,1)$-labeling problem on graphs. *SIAM Journal of Discrete Mathematics* 9: 309–316.

Hale, W.K. 1980. Frequency assignment: Theory and applications. *Proceedings of the IEEE* 68: 1497–1514.

van den Heuvel, J., R.A. Leese, and M.A. Shepherd. 1998. Graph-labeling and radio channel assignment. *Journal of Graph Theory* 29: 263–283.

van den Heuvel, J., and S. McGuinness. 1999. Colouring the square of a planar graph. Preprint.

Hurley, S. n.d. Frequency assignment research. `http://www.cs.cf.ac.uk/user/Steve.Hurley/freq.htm`. Accessed August, 2000. Software description included.

Leese, R.A. 1997. A unified approach to the assignment of radio channels on a regular hexagonal grid. *IEEE Transactions on Vehicular Technology* 46: 968–980.

________. 1999. The mathematics of radio spectrum management. Research program report at `http://www.radio.gov.uk/busunit/research/mathsp/mathsp.htm`.

________. 2000. A linear programming approach to radio channel assignment in heavily loaded, evolving networks. Preprint.

MacDonald, V.H. 1979. Advanced mobile phone service: The cellular concept. *Bell System Technical Journal* 58: 968–980.

McDiarmid, C., and B. Reed. 1997. Channel assignment and weighted coloring. Preprint.

Roberts, F.S. 1987. Personal communication.

Shepherd, Mark. 1998. Radio channel assignment. Ph.D. thesis, Merton College, Oxford University. `http://www.maths.ox.ac.uk/combinatorics/thesis.html`.

Smith, D.H., and S. Hurley. 1997. Bounds for the frequency assignment problem. *Discrete Mathematics* 167/168: 571–582.

Acknowledgment

The author's research was supported in part by NSF grant DMS–0072187.

About the Author

Jerry Griggs is a graduate of Pomona College and MIT, where he earned his Ph.D. in 1977. Since 1981, he has been at the University of South Carolina, where he is Professor of Mathematics and a member of the Industrial Mathematics Institute. He received the 1999 award at the University for research in science and engineering.

His research area is combinatorics and graph theory, both fundamental theory and applications to database security, communications, and biology. He has published more than 60 papers and supervised 11 doctoral and 9 master's students. He serves on the Board of the Mathematics Foundation of America, which oversees the Canada/USA Mathcamp. He has been an MCM judge since 1988.

Space Aliens Land, Threaten Global Destruction

Detroit: “Little green men” land and complain about noise

Space aliens landed simultaneously in all of the world’s major cities at about 3:25 EST this morning. Citing excessive radio noise on Earth’s part, they proceeded to spraypaint a hexagonal lattice onto the cities from low-flying spaceships. A spokesbeing for the aliens then demanded that all of Earth’s radio transmitters be relocated to the center of one of the hexagons and retuned by one month from today. If the transmitters are not relocated and retuned by then, the spokesbeing threatened to destroy the Earth.

In response to widespread human protest, the spokesbeing, who said his name was “Jymyzzach,” which loosely translates to “Jared the Terrible” in English, defended the aliens’ actions.

“Listen,” he said “You humans are using many times more bandwidth than you need and you still manage to have interference and bad reception in some places. My people are astronomers: We survey the furthest reaches of the cosmos for clues as to the secrets of the universe. Whenever we look at the side of our sky that contains your planet, all we can see at all radio frequencies is this huge, brilliant ball of noise, noise, and more noise.”

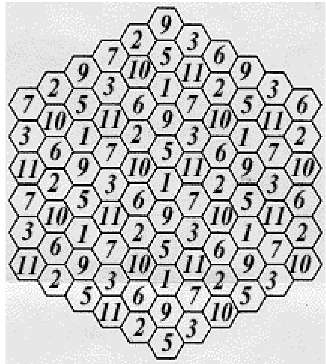
“And your taste in music is deplorable,” added Jared, displaying a false-color radio image of “The Macarena.”

Jared explains the need for assigning frequencies as follows: “Basically, what we decided was that to reduce the noise from your planet, you need to keep the range of frequencies of your channels to a minimum. With the hexagonal lattice proposed, you can cover your cities with signals that do not interfere. So you’ll have clear, crisp signals, and you’ll hear what you want to hear. We even accounted for your inferior technology in our calculations, because when your Earth-transmitters are too close together, an awful interference occurs that we cannot stand!”

“You Earthlings will get better radio reception than ever using no more than 15 channels, and my people will be able to continue our quest for knowledge. We have a win-win situation here,” Jared proclaimed.

Jared also has plans for rural and other areas that the aliens have not divided into hexagons. “It’s hard to make up a small set of rules for maximum bandwidth reduction when radio transmitters are randomly distributed, but with some thought and some computation you can drastically reduce the bandwidth used by these scattered transmitters.”

The UMAP Journal 21 (3) (2000) 387–388.

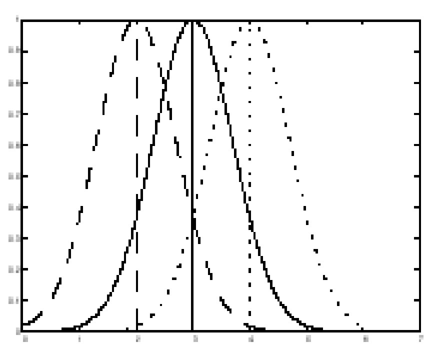


The aliens' transmitter restructuring plan.

When pressed for details, Jared explained his scheme for the unpartitioned areas. "You start by assigning each transmitter a number. You set transmitter number one to channel one. Then you set transmitter number two to channel one and see if it interferes with transmitter number one. If it does, you set transmitter two to channel two and check again for interference; but if it doesn't, you leave transmitter two at channel one and move on to transmitter three. By repeating this process until all transmitters have channel numbers assigned so that none of them interfere, you can get good coverage at pretty low bandwidth."

"When I say that two transmitters interfere, I mean that they can both be heard clearly on the same channel on a radio somewhere. If you draw circles around two transmitters set on the same channel to mark the places where they're just barely audible, you can tell whether they interfere or not by whether those circles intersect. Unfortunately," he continued, "transmitters don't broadcast just on the channel they're set on. They also broadcast a little bit on every other channel, so two transmitters can interfere even if they're set on different channels. You have to take that effect, called 'spectral spread,' into account when you're looking for interference."

Before departing, Jared ordered one of the alien spaceships to destroy the moon. "Just to let you know we're serious," he explained. President Clinton could not be reached for comment.



Different transmission channels interfering.

— Christopher R.H. Hanusa, Anand Patil, and Otto Cortez, in Claremont, Calif.

Elephant Population: A Linear Model

Nathan Cappallo
Daniel Osborn
Timothy Prescott
Harvey Mudd College
Claremont, CA

Advisor: Michael Moody

Introduction

We use a matrix to model the effects of the darting on the population. Assuming that the age distribution was stable at the inception of darting, we:

- drop the birth rate by a chosen factor to simulate a percentage of the elephant cows being darted;
- manipulate this factor to model the waning effectiveness over time of the contraceptive used, thus obtaining an accurate estimate of how many cows to dart each year;
- assume the cost of darting to be comparable to current elephant contraceptives and compare this to the cost of removal; and
- model the effect of darting on populations drastically reduced following a disaster.

The resulting algorithm is sufficiently simple and fast and could be used by many different elephant parks.

Assuming that culling is not a viable alternative, removal appears to be a more effective solution, since darted elephants will need to be darted multiple times over their lifetime. However, this result does not take into account the increasing cost of removing elephants as humans encroach on their habitat. Also, since the relative number of older, bigger elephants will be greater, tourists revenue will increase. Thus, while removal is less expensive, it may not be the best alternative.

The UMAP Journal 21 (3) (2000) 389–402.

Elephant Populations

Growth Rates for Elephant Populations

We first find the rate at which elephants can produce female offspring. Because of the assumed parity between males and females, we ignore the males and assume that the male population is equal to the female population. We feel that this is a safe assumption because the population growth is proportional to the number of females, not to the total number of animals.

Given that elephants produce offspring every 3.5 years, this gives us a starting rate of 1 elephant/3.5 years. Because 1.35% of births are twins, the birth rate is 1.0135 elephants/3.5 years. Finally, only half of the births are females, which gives the final birth rate of $\frac{1}{2} \times 1.0135$ female elephants/3.5 years, or an average of 0.1448 females born per cow per year. This is a good approximation because of the large number of elephants in the herd; any random variation tends to cancel itself out.

Given the unusually long gestation period, a further formula is needed to find the average birth rate per female during her first two years of maturity. Because elephants first conceive between the ages of 10 and 12, we assume that half conceive in the first year and the other half conceive in the second. Since the gestation period is not an integral number of years, one-twelfth of the elephants give birth during the ages of 11 to 12 (half of the elephants were able to conceive, one-sixth conceived during the first two months, so that 22 months later one-twelfth gave birth). The remaining five-sixths of the elephants that conceived during age 10 give birth when they are 12. Also, another one-sixth conceive during the first two months of their eleventh year, so they give birth at the end of their twelfth year. This means that $\frac{1}{6} + \frac{1}{2} \times \frac{5}{6} = \frac{7}{12}$ give birth in their eleventh year. From then on, we assume that elephants give birth yearly, always to 0.1448 cows each birth. This noninteger value reflects that we are dealing with averages, not individual cows.

It is a common practice in biological studies to view the age structure as a vector and the survival and birth rates as an appropriate square matrix. For the population vector, the ith element represents the current population that is between $i - 1$ and i years old. Multiplying the matrix by the vector gives an equal-size vector that represents the population in the next year.

The example below shows the age structure of a species that lives to the age of 8 years. Note that there is a 75% survival rate for the first year, after which the survival rate for any one year is β, except for the next to last year of the animal's life, when the survival rate is $\beta/2$. In the second year of the animal's life, it produces an average of 0.097 offspring. In the third through sixth years of the animal's life, it produces an average of 0.290 offspring per year; during years seven and eight, no offspring are produced. Next to the matrix, the age structure vector gives the relative ratios of the animal ages.

The first row consists of the birth rates. The product of this matrix with a population vector represents the passage of one year. Thus, multiplying the

$$A = \begin{bmatrix} 0 & 0 & 0.097 & 0.290 & 0.290 & 0.290 & 0.290 & 0 & 0 \\ .75 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta/2 & 0 \end{bmatrix}; \quad \vec{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix}$$

matrix by a vector representing the population in one year, you get a matrix whose value in the first entry is the sum of all the female elephants, scaled by the number of births they had; this is the number of newborn elephants the following year.

By adding survival rates in the diagonal directly beneath the main diagonal, the entries other than the first become the population of the year before, scaled by this survival rate. All of the other entries are zero, since elephants can age only one year from the year before and can give birth only to newborns.

We assume that the age ratios have reached an equilibrium; this is safe to say, provided that park management did not selectively hunt or relocate elephants based on age. Then the age vector of the next year is proportional to the age vector of the current year. In other words, $A\vec{x} = \lambda\vec{x}$, in which λ and $\vec{x}$ are an *eigenvalue* and an *eigenvector*. In this instance, λ gives the growth rate and $\vec{x}$ gives the age distribution of the current population, which can be scaled appropriately to fit the known population size.

Given a survival rate of elephants for their first year of between 70% and 80%, we set the first-year survival rate at 75%. The average elephant lives to 60, and the growth rate is 5% yearly [Douglas-Hamilton 2000]. This length of life leads to a survival rate is approximately 99% from one year to the next after the first year.

Because elephants do not live to 70, a lower survival rate is required during their last few years—a decrease from 99% at 60 to 0% at 70. Because these are average values, the survival function should be smooth; it should also be level during the first few years and then decrease more rapidly, much like a sine curve. We use

$$S(a) = 0.99 \sin\left(\frac{(a-48)\pi}{22}\right),$$

where S is the survival rate of an elephant a years old. While this curve may be a theoretical construct, any realistic death rate would have to be similar to this curve and thus have a similar effect on the rest of our math.

Age Structure and Growth Rate

We used a mathematical solver to investigate our 70×70 square matrix (which we do not display here); it has $\lambda = 1.043$. This means that the elephant population grows 4.3% each year. From the eigenvector, we found the age distribution of the elephants (**Figure 1**).

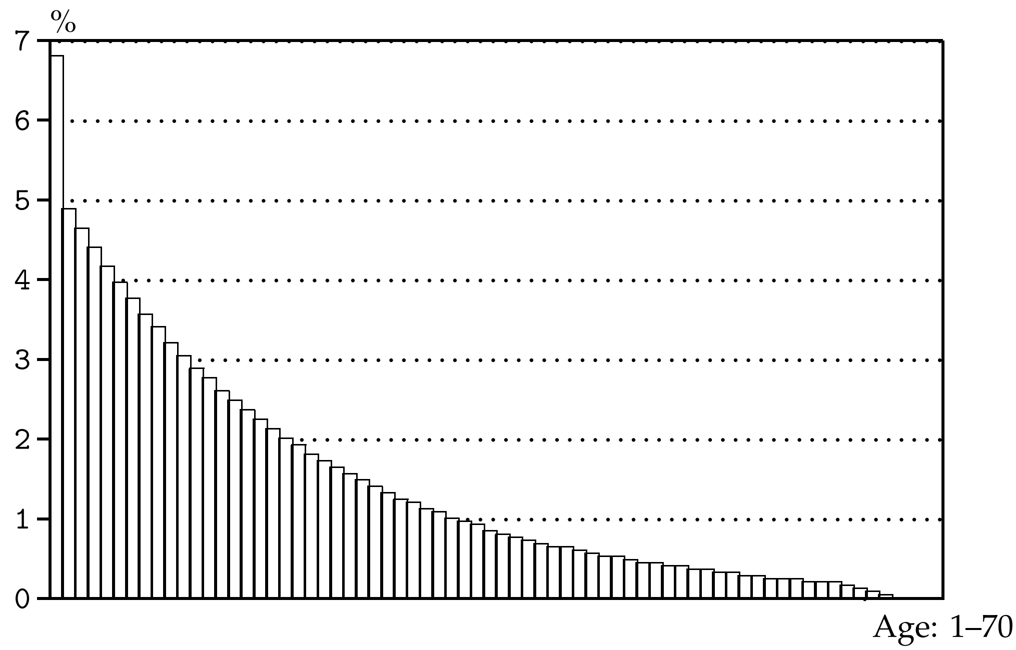


Figure 1. Current elephant age structure.

This growth rate does not agree with the data for removal of elephants from the park in the last two years (20% removed per year). The removal data do not take into account further elephants who may have been culled, so the numbers may have been even higher.

A second source [`www.africalibrary.org` 1999] confirms our estimates of birth and survival rates. Putting maximal values into the matrix, we could not obtain a growth rate of 20%; even with elephants giving birth every 22 months and a death rate of zero, the growth rate was never above 15%. So we decided that the data on removal were erroneous.

Darting Elephants—Now and Tomorrow

How Many and Which Ones?

To maintain a zero-growth population, we need to reduce the birth rate to 27.1% of its current value μ. This is a much more drastic change than the removal of the 600–800 animals that was required over the past 20 years. The main reason is that the birth rates are already low, so a much larger change is needed to affect the population. The survival rates, on the other hand, are high, so a much smaller reduction can affect a larger change.

We assume that the drug is 99% effective immediately upon injection and is still over 90% effective at the end of the first year; by the end of the second year, the drug drops to zero efficacy. A particular percentage of efficacy means that that percentage of the population is still under the effect of the drug and cannot conceive. Like the sine wave for survival rates, the drug efficacy should be concave downward, decreasing more rapidly as the second year ends. However, we feel that the sine wave decreases too rapidly for the first few months and too slowly in the end to model correctly the effects of the drug. We use instead a fourth-order polynomial:

$$E(t) = -.062t^4 + .99,$$

where E is the drug's effectiveness at time t years after injection.

This means that there is an average of 97.8% efficacy over the first year (implying a 2.2% chance of pregnancy) and 60.6% efficacy over the second year. This also means that a cow darted both years would have a 0.9% chance of pregnancy. Denote by γ the percentage of elephants darted; then in two years' time, the percentage darted both years is γ^2, the percentage darted only once is $2\gamma(1-\gamma)$, and the percentage not darted either year is $(1-\gamma)^2$. This means that the percentage able to get pregnant (and hence the factor by which the birth rate drops) is

$$\mu = 0.00883\gamma^2 + (0.3944 + 0.0224)\gamma(1-\gamma) + (1-\gamma)^2,$$

which simplifies to

$$\mu = 0.59203\gamma^2 - 1.5832\gamma + 1. \tag{1}$$

Setting this expression equal to the desired birth rate reduction (to 27.1% of the current value) gives the desired darting rate γ as 59% of all reproductive female elephants per year.

We also need to find the targets for the contraceptive. Should the park staff seek to drug whole herds? a certain percentage of each herd? or every animal in a certain age group? While seeking whole herds would be most cost-effective, this practice would decimate the herd by not allowing it to reproduce. Seeking out specific ages would be too expensive because of the difficulty in determining the ages of specific elephants. Therefore, we decided that we would target a random percentage based on the value of γ.

With the birth rates reduced by μ, the growth rate is 0.0%, as desired, and the stable age structure (achieved some time in the future) is shown in **Figure 2**.

Uncertainty in Derived Data

To find how the uncertainty in our given data affects our estimate, we could propagate the uncertainty through the functions involved. But error propagation through the process of finding the eigenvalue of a 70×70 matrix is tedious,

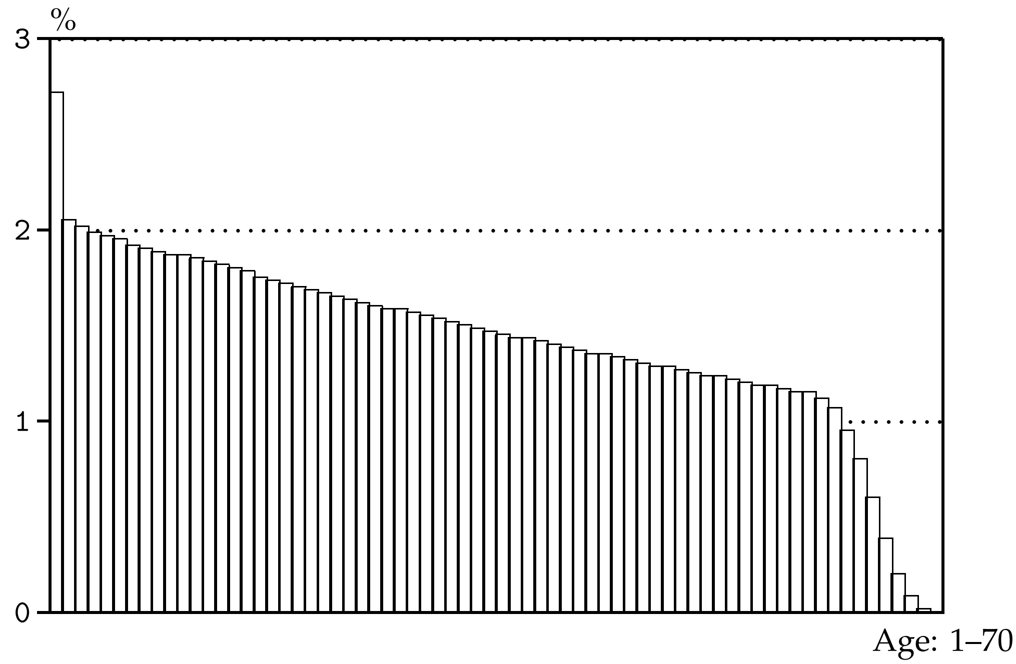


Figure 2. Equilibrium age structure with darting 59%.

so we use another method. We put into the population matrix all of the values that would cause a higher-than-calculated birth rate and use this to find the resulting error in μ.

First, we take the maximal scenario: 100% survival for adults, 80% for juveniles. Then the birth rate must be reduced by 82%, 9% more than the calculated value of 73%.

Taking the minimal scenario, we encounter a problem. We take the survival rate to be 99%, because anything lower causes the percentage growth to become unrealistically small. We solve this by taking the error in the growth rate of 4.29% to be $\pm 1\%$. This allows values as high as 5.29%, which is in agreement with the maximal scenario's pre-darting eigenvalue, and implies a growth rate of 3.29% for the minimal scenario. Assuming the juvenile survival rate drops to 70%, this implies a general survival rate of 0.984. Upon plugging these values into the matrix, we find that we need to reduce the birthrate μ only by 64%, again 9% away from the calculated 73%.

So, our estimates point to an error no larger than $\pm 9\%$.

Next, we want to find the error in how many elephant cows we need to dart. Our final goal is to estimate the error in the costs of darting and removing elephants.

Taking the derivative of both sides of **(1)**, we get:

$$\partial\mu = 0.59203 * 2\gamma * \partial\gamma - 1.5832 * \partial\gamma.$$

Solving for $\partial\gamma$ gives

$$\partial\gamma = \frac{\partial\mu}{0.59203 * 2\gamma - 1.5832}.$$

Taking the calculated value for γ to be 59% yields a value for $\partial\gamma$ of 0.0595, giving us the second step in our process: the uncertainty in how much we need to dart is $0.59 \times 0.0595 = 0.035$; therefore we need to dart $59 \pm 3.5\%$ of the elephants.

Sensitivity and Stability

The rather high uncertainties—3.5% for the percentage to dart and 15% for the cost (derived later)—point to low general stability.

Upon changing individual values in the population matrix, we find that the value for the survival rate is the most important. Changing that rate by only 1%, the growth rate changes by up to 1% and the number we need to dart by 9%. Birth rate and juvenile survival rate are insignificant by comparison, as was the gestation time.

Age Structures of the Future

What will the elephant population look like in 30 years? Assuming that the population was already at equilibrium without darting, we multiply the age vector by the new, adjusted matrix, yielding a vector that represent what the age structure would look like in the next year. Reiterating another 29 times, we find what the age structure would look like in 30 years (**Figure 3**).

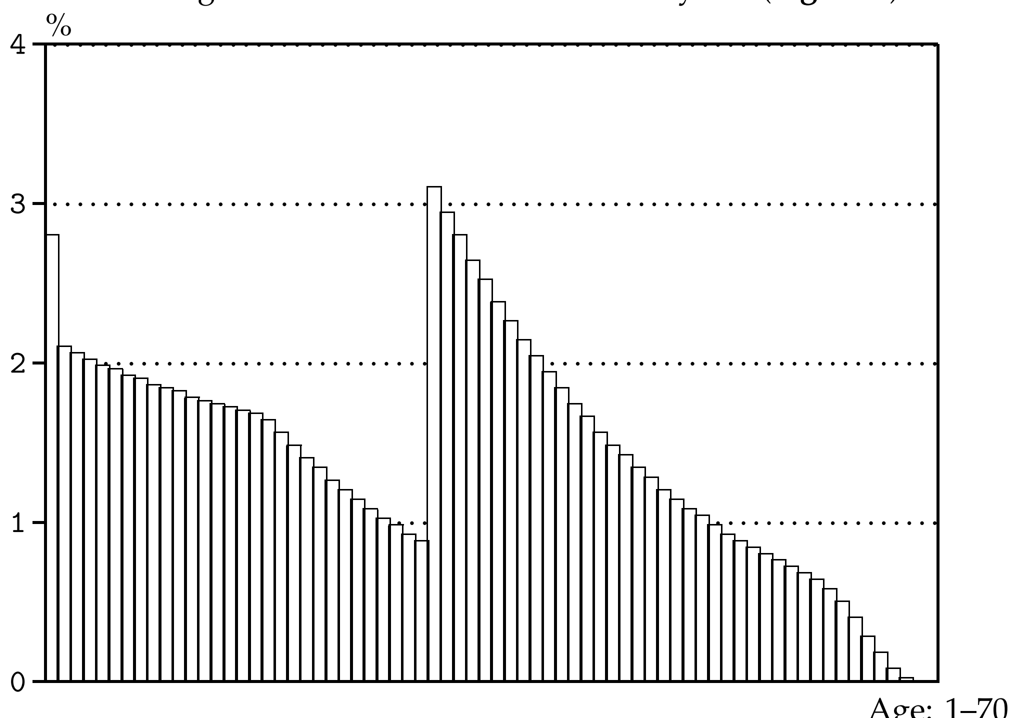


Figure 3. Age structure in 30 years.

The most notable aspect is the sharp peak that occurs at 30 years. At first glance, this may seem troubling because of the unexpected discontinuity; however, this is to be expected, because it represents the sharp drop in births when the contraceptive program began 30–31 years earlier.

Another interesting detail is the spike in the zero-year-old range. This is a result of many of these elephants dying off before they reach age one, as reflected in the value of 0.75 for their survival rate.

Repeating another 30 times allowed us to find the age structure 60 years after the darting process (**Figure 4**).

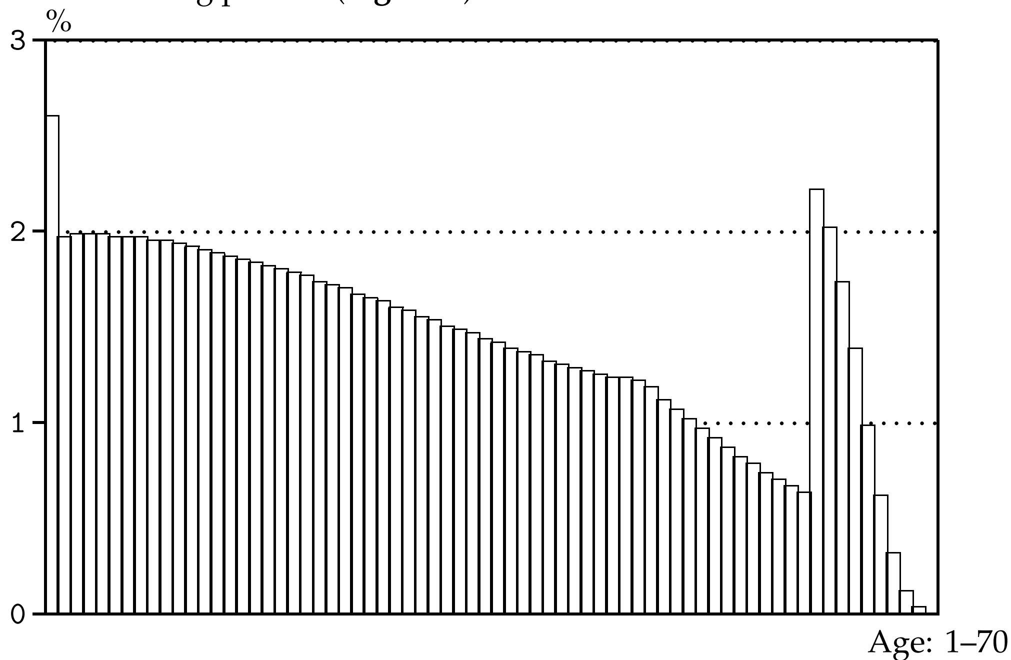


Figure 4. Age structure in 60 years.

The aspect of this situation that is most troubling at first is the large jump in the population structure in the elderly age range. However, this is to be expected once again, as it points to the last year that the elephants were not darted. A more interesting aspect of the graph is the slight increase of the two- and three-year-olds as opposed to one-year-olds. The reason is that the elderly group that we were examining a moment ago contributed to this slight increase; then they were no longer able to contribute to the current one-year-old population because they had reached the age of 61, at which point elephants no longer bear young. The population dip at the one-year mark is a result of this effect.

Given the age structure some number of years after the darting began, we find the difference between that year's expected age structure and the calculated equilibrium age structure. After weighting each age vector to have the same total population (their sum), we find the difference between the two and calculate the length of the difference vector. Plotting this length over time, we were able to find the difference from any age vector to the expected equilibrium value (**Figure 5**).

A large hump occurs at about 60 years. This happens because up to this point the large spike that was prevalent in the age distributions at 30 and 60 years prevents the age vector from getting closer to the equilibrium vector. At about 60 years into the future, however, this large spike begins to die off as the elephants in that age group reach age 60 and begin to leave the population.

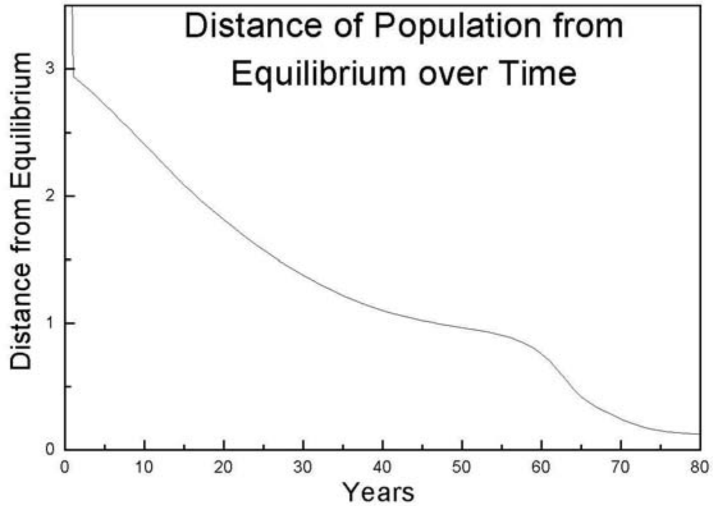


Figure 5. Difference between projected age vector and projected equilibrium value.

Effects of Darting on Tourism

Contraceptive darting is conducive to increased tourism, not simply because it means an end to killing but also because of the changes it creates in the age structure. By shifting the population distribution to favor older elephants and keeping the total population size the same, darting assures a higher number of older elephants in the future. Older elephants are not only bigger, but smarter; a herd led by an experienced elephant is much more stable and tourist-friendly than one led by an inexperienced elephant [Mullins 1997].

Removal of elephants, while cheaper, may not be best when one takes tourism into account. The increased tourism due to bigger, calmer elephants may bring in revenues that far exceed the extra spending. Here a model cannot help us, only time can tell which method is cost-effective.

Relocation vs. Darting

Our model shows that it is possible to rely completely on darting female elephants to control the growth rate and keep a stable population, by darting roughly 59% of the fertile cows yearly.

If we simply remove elephants each year, we are paying a flat cost to remove them from the population forever; removal can be modeled by increasing the death rate of the elephants. When we dart elephants, however, we are not removing them, so the population still increases unless we reduce the birth rate to be as low as the death rate. We also have to re-dart the females every year to keep this birth rate lower, meaning that we are darting more females than the total number of elephants that we would be transporting. Depending on the costs of darting, this could become a more expensive endeavor than simply removing several elephants every year.

It costs \$800 per elephant [Shaw 1999] to move an elephant out of the park, and the cost of darting a single wild horse with the same contraceptive is \$25 [Bama 1998].

How much would it would cost to dart an elephant with this contraceptive? The contraceptive works by stimulating the immune system of the mammal to produce antibodies that bind to the sperm receptor sites of the oocytes [`www.wildnetafrica.com`]. This means that when the sperm attempt to bind to the eggs of the mammal, there are no places for them to bind to, and hence the sperm do not fertilize the egg. The immune system needs to be triggered by a sufficient concentration of the antigen, so the dosage should depend on the total mass of the animal in question. With this information, and the fact that the average mass of an elephant (3250 kg [Estes n.d.]) is 9.7 times the average mass of a wild horse (335 kg [`www.agric.nsw.gov.au`]), it should take 9.7 times as much contraceptive to have the same likelihood and strength of a result. At \$25 per horse, the cost for an elephant should be about \$242.53, or after costs of the helicopter fuel and maintenance, about \$250.

We want to determine the total costs for various levels of removal and darting. The difficulty is that the more elephants we remove from the park, the fewer elephants we need to dart, and we cannot easily tell how one changes with the other. Random darting can be simulated by scaling all the birth rates down by a fixed amount and removal by a change in the death rate. Both rates are measures of how the population, and more specifically the age groups, change from one year to the next. If we assume that the rangers remove animals regardless of age, then we can create a fixed value for the removal rate and multiply it times all of the death rates.

The two variables that we can control are the darting ratio γ and the removal number ρ. The proportion not removed in a given year is $\sigma = 1 - \rho/C$, where C is the (varying) total number of fertile cows in the population.

We test two extreme cases, $\rho = 0$ and $\rho = 300$, to see how they affect C. We tested $\rho = 300$ under the assumption that C does not change from year to year. This results in a variation in C of 2% and the same in σ. This tells us that both C and σ can be treated as constants.

We alter our original model population matrix by multiplying all of the survival values (that is, the values in the diagonal directly below the main diagonal) by σ. We choose a value for ρ and get a value for σ from it. We place this in our matrix accordingly, and then test the positive real eigenvalue. If it is greater than one, then we try a value of μ and manipulate it until we achieve $\lambda = 1$. At this point we have a stable equilibrium value. We repeat this process several times for the different values for ρ and recorded.

The last step is to find the costs. We solve for γ and multiply by \$250 per cow to get the cost of darting. We get the cost of removal by multiplying the ρ value by \$800 per cow. The sum of these values is the total cost, shown in **Figure 6**.

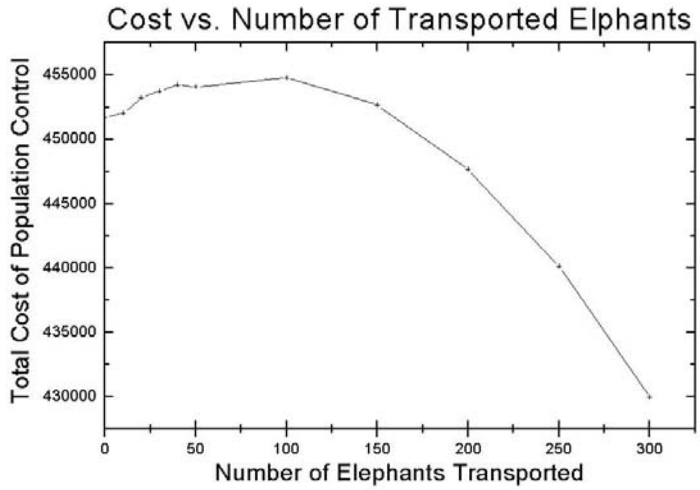


Figure 6. Cost of control as a function of number of elephants removed.

Recovery from Drastic Loss

Our model takes advantage of the fact that for large populations, fluctuations tend to cancel each other out. That is why the birth and death rates are so constant. For example, if 50 cows die due to an epidemic of flaccid trunk disease, out of a population of 100 this is a catastrophic 50% loss! However, if the population is 10,000, this is merely 0.5%. Thus, a variation that is a disaster for a small population is negligible for a larger one.

Therefore, for modeling the effect on the population due to a catastrophic loss, a different model is needed, since the assumption that birth and death rates are constant is no longer valid. This model must be flexible enough to account for differing death rates for a given year.

Since the only thing we need to know is how the darting affects the outcome, we can use marginal analysis. Thus we can start with a loss of 90% and see how much of a death rate it would take for the population to die out.

A certain minimum number of a given species is needed for viability. For this minimum value, we choose 20 elephants, about the size of a single herd.

Applying this analysis to our specific case, if 90% of the 11,000 elephants die off, we have only 1,100 left. To address the effect that darting might have on the situation, we consider two different cases: if the elephants are being darted, and if they are not.

No Darting

For initial population p_0, the population one year later is $p_1 = p_0 + p_0(b-d) = p_0(1+b-d)$, where b is the birth rate and d is the death rate. For the sake of simplicity, we set the birth rate to be the average value of 0.1448, assuming that any change will be illustrated by an appropriate change in the death rate.

Then, doing algebra to turn the recursive function into a general one, we find $p_n = p_0(1 + b - d)^n$.

Solving for d, given $b = 0.1448$, $p_0 = 1{,}100$, $p_n = 20$, and $n = 10$ years, we find (after discarding the unrealistic solution) that for the elephants to be doomed in 10 years after a 90% loss, d would have to be 47% on average! This is an extraordinarily high number, especially when one notices that this was the worst-case scenario.

Darting

For a worst-case scenario, let the first 5 years have a birth rate of 0. After that, the rest of the $n - 5$ years have a normal birth rate, taken again as 0.1448. Thus one gets the general equation $p_n = p_0(1 + b - d)^{n-5}(1 - d)^5$.

Taking $n = 10$ again, $p_n = 20$, and $p_0 = 1{,}100$, we find that the 10-year "doom" value for d becomes 40%.

Conclusion

We conclude that the effect on survival of darting ($d = .40$ vs. $d = .47$) is minimal.

Other Park Populations

The matrix method that we have applied to the herd can be used with other populations, of elephants or nonelephants. Difficulties arise only if the park is too small, when small variations in data might be amplified and the model possibly lose stability and accuracy.

Report to Park Management

While at first glance modeling may seem quite abstract, in reality it is an extremely useful and practical tool. The universe contains patterns that may be studied and used to make predictions. These patterns are everywhere: human beings need to eat and breathe, a thrown ball will eventually return to the ground, ice cubes never form spontaneously in a cup of coffee. It is this underlying order that a model takes advantage of to make predictions about the future.

For all the complexity in elephant populations, there are many patterns to it. Most of these are simply common sense. Take the elephant gestation period, for example; it can be used to make an estimate of the maximum elephant fertility.

Other patterns are the nearly constant survival rate after age 1, the constant birth rate, and negligible migration. They allow us to make predictions about the behavior over the elephant population as a whole.

It is easier to model the behavior of the whole preserve than that of an individual elephant. For larger, more complex systems, patterns become even easier to see and understand, because random variations cancel each other out. There is no such thing as a statistical certainty; but the larger a group, the fewer fluctuations due to random chance, and the surer one can be.

We took advantage of these facts in making our model of the elephant populations in your park. Random chance effects tend to cancel each other out in so large a population. Thus, we are able to use a uniform and constant value for the birth and death rates of the elephants.

Given these rates, we modeled the changes in the population over time by putting them into a matrix. This matrix, when multiplied with a list of elephant populations for every age in a given year, returns a list of how many elephants would be that age the next year.

We assume that the age distribution of the elephants has become constant over time. In essence, this says that as elephants get older, younger elephants also get older and replace them. Some of the younger ones die before they can get older, however, so the population numbers remain the same.

Without culling, removal, or darting, the population would tend to grow but the proportion of each age would remain the same.

Such details as twins, gestation, and gradually increasing death rates can also be accounted for. These details can be simulated by slightly altering a few of the numbers in the matrix.

Once we can model the population, simulating darting becomes simple. We make the rather safe assumption that darting the elephants drops the birth rate to a uniform degree. We find that the growth rate of the elephants needs to be reduced to 27% of the natural reproductive rate. To accomplish all of the reduction through darting, 59% of females would need to be darted each year.

But the darting will also affect the age structure of the elephants, and our model easily simulates this over any number of years.

However, for smaller systems, our approximations of constant, average values are no longer valid. This is not to say that we cannot model smaller systems, only that the larger the system, the more effective a model can be.

We modeled a much smaller elephant population subjected to catastrophic loss. Take the worst disaster you can think of—say 90% of the elephants dying—and then say that somehow the death rate experiences a continual "random fluctuation" upward constantly, for 10 years. Even with darting it would take a constant death rate of 40% per year to kill off the elephants. Hence, use of contraceptive darts would have no effect upon whether or not the unlikely event of a sudden loss would cause an extinction of the park's elephant population.

The data for the last two years are startling: 20% of the elephants were removed each year! This seems to imply that the population is growing at 20% per year. However, such a rate is unrealistic.

Much of our confidence in our model comes from testable predictions that it makes. When we plot the age structure 30 years after darting, we see a graph that gradually dies down until it reaches the number of 30-year-olds, where there is a sudden spike, corresponding to elephants born before darting. Furthermore, the graph of the 60-year-olds shows not only this spike but another bump about 10 years later, reflecting the formerly larger number of elephants having given birth to a larger number of calves. These effects are predicted by the model without any manipulation by us.

With common sense, a little data, and some mathematics, an excellent predictive model can be made. Our model should be a useful tool in your making well-informed decisions.

References

Bama, Lynne. 1998. Wild horses: Do they belong in the West? *High Country News* – March 02, 1998. `http://www.hcn.org/1998/mar02/dir/Feature_Wild_horse.html`.

Douglas-Hamilton, Iain. 2000. Discovery Online, Field Notebook: Living with Elephants — Q & A. `http://izzy.online.discovery.com/area/nature/elephants/qanda.html`

Estes, Richard. n.d. Elephant. `http://www.nature-wildlife.com/eletxt.htm`.

Mullins, Lisa. 1997. Living on Earth Transcript: January 1, 1997. `http://www.loe.org/archives/970103.htm#Delinquents`.

Shaw, Angus. 1999. Zimbabwe moves elephants out of harm's way. `http://www.canoe.ca/CNEWSFeatures9904/06_elephants.html`.

`http://www.africalibrary.org/course/papers99/yankoupematamala.html`. 1999. Page no longer available at this URL.

`http://www.agric.nsw.gov.au/mdil/horses/ax24.htm`. Page no longer available at this URL.

`http://www.wildnetafrica.com/wildlifeorgs/oprepro/#IMMUNOCONTRACEPTION`. Page no longer available at this URL.

A Computational Solution for Elephant Overpopulation

Jesse Crossen
Aaron Hertz
Danny Morano
North Carolina School of Science and Mathematics
Durham, NC

Advisors: Dot Doyle and Dan Teague

Introduction

We extrapolate longevity data and explore the long-term behavior of the population age distribution. We determine the number of dartings to fix the long-run stable population at 11,000; about 1,300 dartings are needed for an every-other-year strategy. We employ two simulations, one based on averages and the other tracking each elephant individually, whose results agree closely.

Our modeled population recovers from sudden declines and is not overly sensitive to small changes in survivorship data. The model also allows estimating the number of dartings if up to 250 elephants are relocated each year.

Assumptions

- We are told that emigration and immigration are rare, so in our model no elephants enter the park except those that are born. None leave except those that die or are relocated.
- Fifty percent of the elephants are female, as the problem suggests.
- It is beneficial to the population as a whole, as well as more economically feasible, to use as few contraceptive darts as possible.
- Cows first conceive when they are 11 years old, rather than some time between ages 10 and 12.
- Gestation always takes 22 months exactly, instead of approximately.

The UMAP Journal 21 (3) (2000) 403–416.

- The darts work, so a cow hit by a dart will not conceive for two years.
- Otherwise, cows give birth every 3.5 years until they reach the age of 60.
- There is a 1.35% chance that a given birth will result in twins.
- The survival rate for the first year is .75.
- The initial population is 11,000 individuals.
- The rangers can readily determine which females are not pregnant, so that no pregnant females are darted, as at Kruger National Park in South Africa, which uses a similar contraceptive program [Purdy 1998].
- Cows normally mate once every 3.5 years. The cycle of a cow darted is not disrupted. If the effect of the dart wears off before she would normally mate and become pregnant, she conceives and gives birth on schedule.
- Previous methods of population control eliminated individuals randomly, so no age group was disproportionately depleted and the relocated elephants have an age distribution that is typical of the population as a whole.
- Since the methods of population control that have been used have no effect on the fertility of the cows, we assume that the initial birth rate is constant.

Analysis of the Problem

We predict the long-term behavior of the elephant population as a function of the number of females. If we track each elephant individually, we must track 11,000 individuals; if instead we look at the population as a whole and take an average-case scenario, we must find formulas for birth and death, mating, aging, and the added effects of the contraceptive darts.

We use both methods. First, we use a computer simulation to track each elephant through its lifespan: We use known probabilities to determine when each elephant is born, reaches maturity, gives birth, and dies. We can use this simulation to test any darting strategy. The results are far less smooth than for an average-case scenario.

We also use another program based on recursive equations to predict the average-case behavior of the population, which we divide into groups of the same age. This method requires far less computer time. The replacement of random events with a deterministic average allows for ready investigation of long-term behavior without interference from individual unlikely events.

Using these two models, we find a mathematical expression for the dynamics of the population and then use these programs to forecast the results of our darting strategy and to demonstrate its stability and flexibility.

Task 1: Predicting Survivorship

There are three distinct phases in the life of an elephant.

- From birth until five years old, the young elephant is very susceptible to predators and accidents and cannot fend for itself while it still nurses from its mother [African Wildlife Foundation 1998].
- After it is weaned, at five years of age, it lives most of the rest of its life in relative safety. Several things can kill an adult elephant, but none has a major effect on the population. There is a low rate of disease, accidents are very rare, and no natural predators can kill something as large as an adult elephant [Hanks 1979, 109]. Therefore, over this period the death rate of the elephant is fairly low, about 2% per year.
- Over the course of its lifetime, the elephant grows six sets of molars; around age 50 the final set of teeth wears out, making it impossible for the elephant to properly chew its food, so that the animal eventually starves to death [Holloway 1994].

We construct a survivorship curve as a piecewise function, with each segment corresponding to one of these phases. Using our assumptions that the given data are a random sample of the elephant population, that the birth rate in the park has been essentially constant, and that the previous killing has been evenly distributed over the population of elephants, we conclude that the demographic shape of this population is typical for an elephant population.

Survivorship l_x is the fraction of the population alive after x years. To compute the survivorship from the data, we sum the data from each year to get a larger sample size and divide the entire data set by the population at age zero. The final survivorship data looks like **Figure 1**.

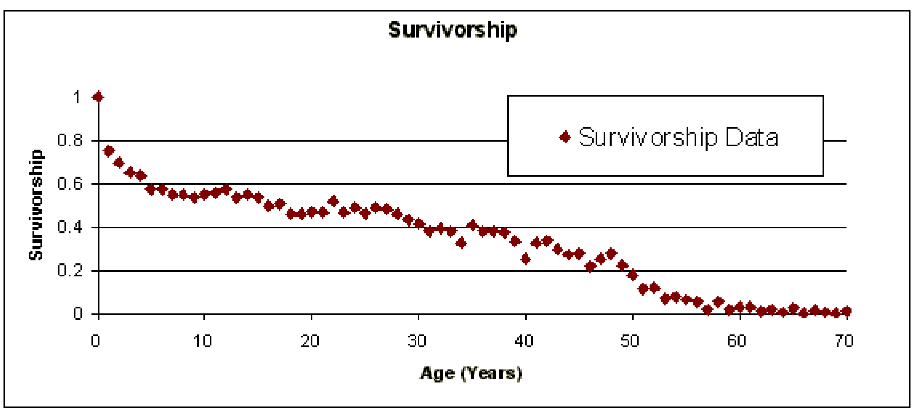


Figure 1. Survivorship function.

The data divide up roughly into three linear sections corresponding to the three stages of the elephant life cycle. These three sections appear to be well approximated by lines, so we generate a piecewise function composed of three linear segments for the ages from 2 to 60, based on a least-squares fit. The two points of discontinuity between the pieces of the function cause little error.

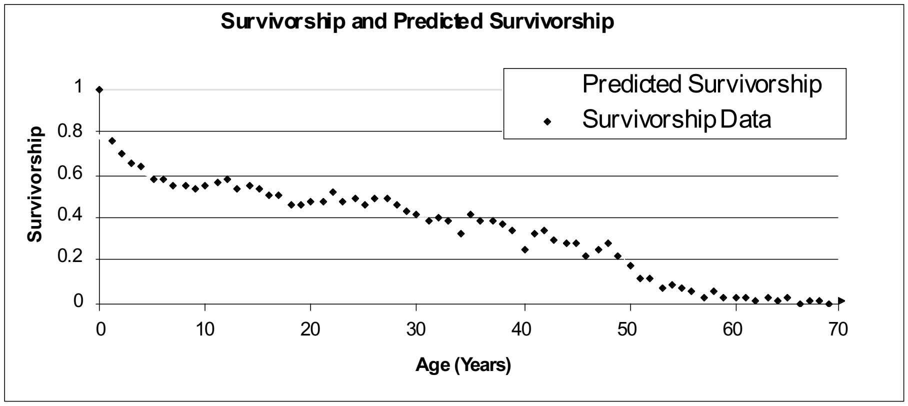


Figure 2. Survivorship data and fitted function.

The survivorship function is

$$l_x = \begin{cases} -0.038806x + 0.77512, & 2 \le x \le 5; \\ -0.007818x + 0.640015, & 5 \le x \le 50; \\ -0.013116x + 0.799663, & 50 \le x \le 60. \end{cases}$$

We calculate the probability of death p_d as the fractional change in l_x:

$$p + d = 1 - \frac{l_{x+1}}{l_x}.$$

The assumption of a constant birth rate is incorrect, as the data are clearly not monotonically decreasing. But given the assumption that previous population control methods (i.e., shooting) did not affect the age distribution, our model is presumably close to the actual profile.

Task 2: Achieving Stability

Birthing cows are females older than 11 and younger than 60 who can give birth; we choose some number D of nonpregnant cows to dart. Because of the additional stress on the darted population and the expense of darting, we should dart as few elephants as necessary.

How often should we dart cows? Darts remain effective for two years. Because the darted elephants are not tagged when they are darted, annual darting would lead to some elephants being darted two years in a row. Darting every two years uses fewer darts and simplifies our solution.

In a population with a stable birth rate, the same distribution occurs among the age groups—each segment of the population has a characteristic percentage.

The only segment that we can directly affect is the fraction f_b that are newborns; the goal is to stabilize the number N of newborns. So, in an ideal setting, after the population stabilizes, we have $N = Ef_b$.

The actual number of newborns is proportional to the number of cows that can give birth in the next year multiplied by the average number of elephants produced at the end of a successful pregnancy and the average chance of a pregnant female surviving long enough to give birth. The average number of elephants born after a pregnancy is one plus the chance p_t of having twins. The average chance of survival $\bar{p_s}$ for up to one year is

$$\bar{p_s} = \frac{\int_0^1 (1-p_d)^t \, dt}{1-0} = \frac{(1-p+d)^1 - (1-p_d)^0}{\ln(1-p_d)} = \frac{-p_d}{\ln(1-p_d)}.$$

The number of pregnant cows that could give birth next year is the number of cows that were not darted two years ago, survived for two years, and are now at least 10 months pregnant. Because cows are distributed randomly throughout the mating cycle, the chance that a pregnant cow is within 12 months of giving birth is 12/42. The chance of a cow having survived for two years is simply $(1-p_d)^2$. The chance that a cow was not darted two years ago is the probability that a nonpregnant cow was not darted two years ago, or one minus the number that were darted two years ago over the number of cows that were not pregnant then. Let P denote the number of pregnant cows. Substituting for the number of cows within a year of giving birth, we find

$$N = (1+p_t)\bar{p_s} \cdot \frac{12}{42} \cdot C(1-p_d)^2 \left(1 - \frac{D}{C-P}\right).$$

We set the real number of newborns equal to the ideal number of newborns and solve for the number of dartings. This tells us the number of elephants that we should have darted two years ago. We base the number of darts to use this year on the effect that the darts had two years ago. Because other terms are constant every year, we can apply the darting equation and find the number of cows to dart this year using this year's C and P. In the case of an excess of newborns, darting increases; if too few births occur then D becomes negative, suggesting that more pregnancies are needed than the population can produce even if no cows are darted.

$$D = (C-P)\left(1 - \frac{11{,}000 f_b}{\frac{12}{42} \cdot C(1+p_t)\bar{p_s}(1-p_d)^2}\right)$$

For values of the parameters, we have $1 + p_t = 1.0135$ and $\bar{p_s}(1-p_d)^2 = 0.94$. The equation should give the number of dartings for tending toward a steady number of newborn elephants How does it behave? To find out, we wrote a program to trace the progress of the population over time. Each year, the number of elephants in one age group times their chance of survival becomes the number in the next age group. We replace the newborn age group with a new generation calculated as the number of pregnant elephants that gave birth in that year times the number of newborns at each birth, $1+p$. **Figure 3** shows the convergent, oscillating pattern that results.

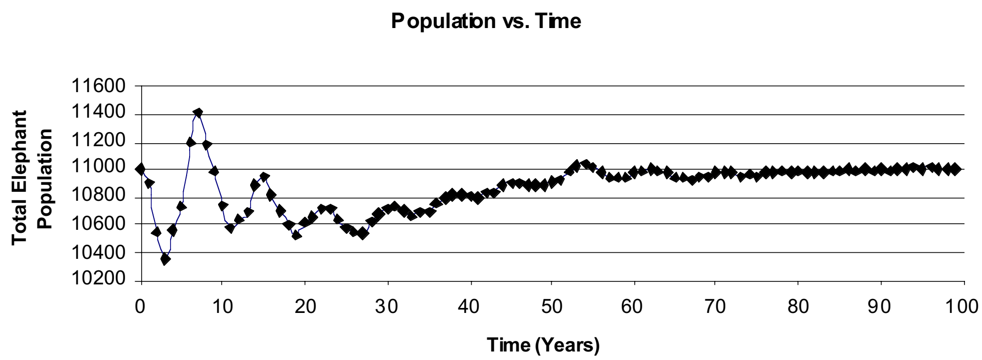


Figure 3. Population over time.

For the first several years after we introduce the contraceptive, the population fluctuates as the model adjusts to compensate by stabilizing the birth rate. In the past, up to 800 elephants were killed every year; here the population never diverges from 11,000 by that much.

How many elephants are darted? While the number initially fluctuates between 0 and 2,000, it levels out to around 1,300 darts per biennial darting, or about 25% of the female population.

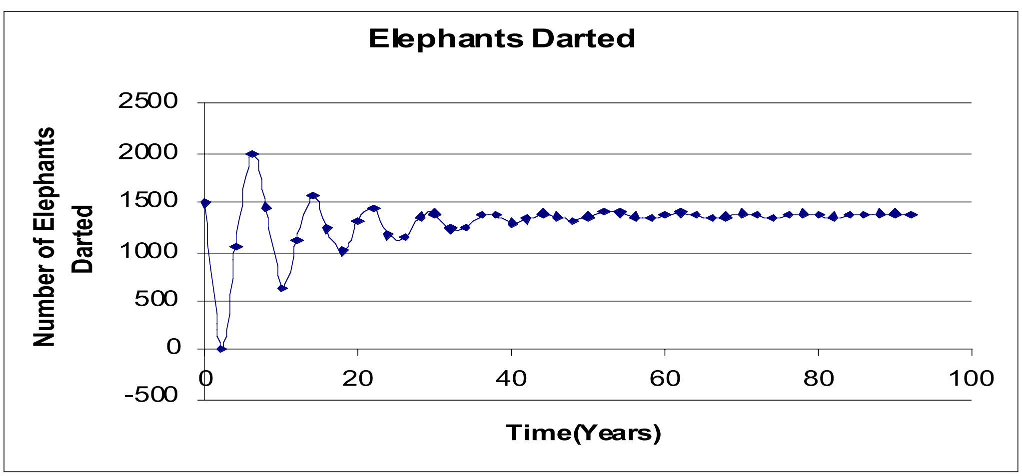


Figure 4. Numbers of elephants darted.

We can simulate the population more accurately by keeping track of each elephant as it ages, gives birth, and dies. Instead of using average probabilities, we use random events to simulate the chaos of the real world. We also keep track of the population on a monthly instead of a yearly basis. **Figure 5** shows a graph produced by our random case simulator. The darting strategy still causes the population converge to 11,000 after some time.

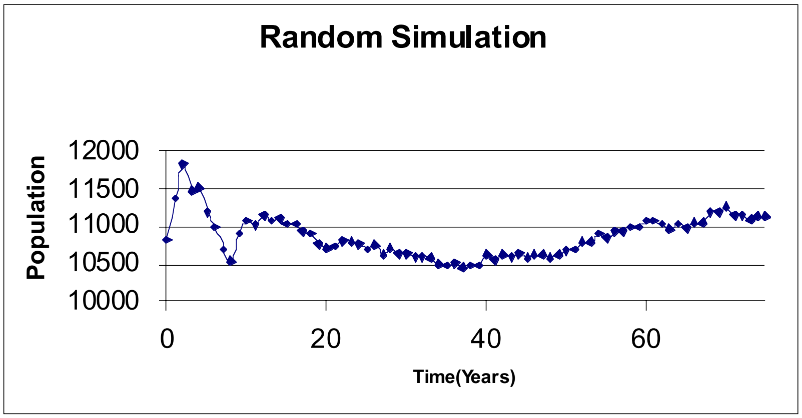


Figure 5. Numbers of elephants darted.

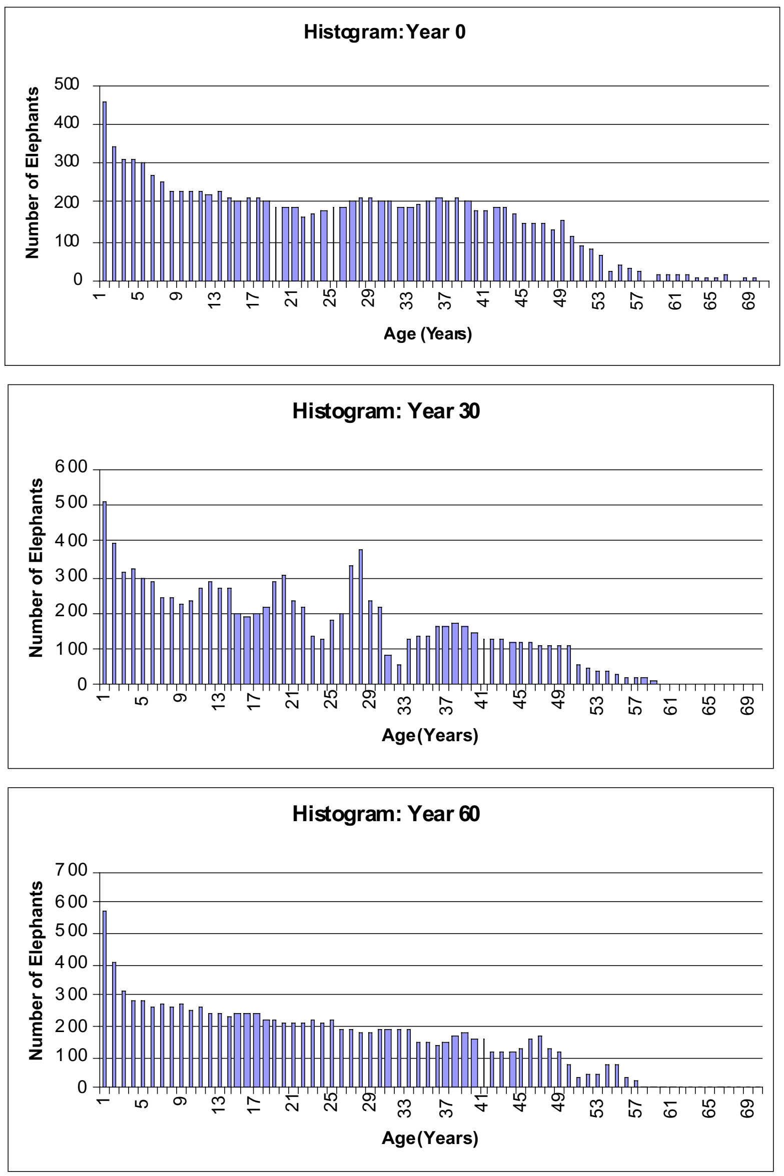


Figure 6. Age distribution initially, after 30 years, and after 60 years of the darting strategy.

Over time, the age distribution tends to shift towards more very young elephants and newborns and fewer old elephants. **Figure 6** shows the initial population distribution and the distributions after 30 years and after 60 years:

- Initially, there is a large number of animals between 25 and 45 and the number of newborn animals is not much larger than all the others.
- After 30 years, there are noticeable spikes in the population due to the large fluctuations that occur during the first several years of the model. There are large numbers of the slightly younger animals, which is good for tourism—tourists usually are attracted to cute animals; additionally, there are still large numbers of the large majestic elephants that everyone wants to see.
- After 60 years, the curve has become much more regular. The only large peak is at the baby elephants. This is the best possible situation for tourists—you can see a good representation of the whole spectrum of young and old, plus a large number of cute babies.

Task 3: Relocation

Relocating elephants each year could make our method more successful, by reducing the number to dart and reducing the stress on females of monthly oestrus. Since we are darting every two years and relocation would remove pregnant and fertile elephants, the combination of darting and relocating has the potential for creating a population disaster; however, we can avoid such a problem by picking the right number of elephants to relocate.

A simulation of relocating 100 elephants per year gives a graph of population much like **Figure 7**.

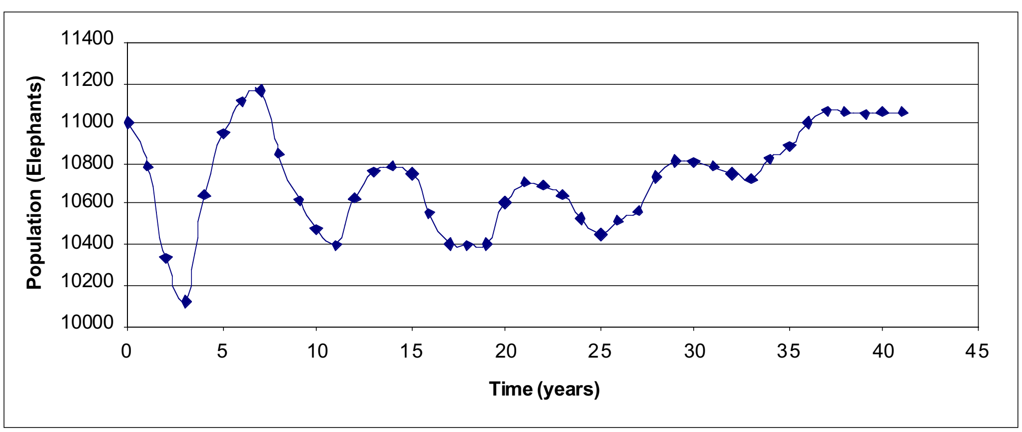


Figure 7. Simulation of removing 100 elephants per year, in addition to darting.

The population drops severely in the first few years but recovers. If this population drop of up to 8% is acceptable, relocation seems to be a viable option. As well as looking at the effects of relocation on population over time, we can also track how many darts we save by relocating various numbers of elephants. The average case results are summarized in **Table 1**.

Table 1.
Darts saved by relocating.

Relocations per year	Darts used in 50 years	Average number of darts per darting	% of darts saved
0	29,900	1,196	0%
50	24,700	988	17%
100	20,200	808	32%
150	16,250	650	46%
200	12,750	510	58%
250	9,700	388	68%

Relocating more than 250 elephants a year could cause an uncontrollable population crash after only a few years.

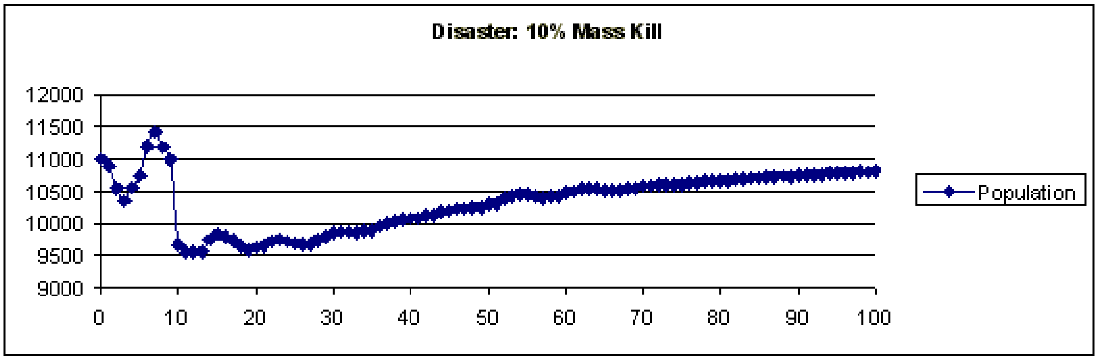
Task 4: Disaster Recovery

Darting may not allow the population to recover from a disaster even if we immediately stop darting. We examine a number of disaster scenarios and see how our model responds to them.

- The first case is a major disaster, such as a rapidly spreading and very deadly disease that indiscriminately kills all segments of the elephant population.
- Next we consider a natural disaster, such as a drought or a famine. In such a disaster, the weakest elephants are most likely to die; these tend to be the youngest and oldest elephants in the population. To model this, we kill portions of the population that are under the age of 10, because they have not yet reached maturity, and portions that are over 50, because they are suffering from the effects of old age.
- Finally, we consider the effect of excessive hunting. Hunters hunt elephants with large tusks, found on very mature elephants. Therefore, we remove parts of the population over the (arbitrary) age of 40.

In each case, we compared removing 10% with removing 50% of the selected population, to simulate moderate and severe disasters. In each scenario, the disaster occurs during year 10.

In the case where 10% of every segment of the population dies, the population hits a minimum of 9,500 and increases fairly steadily thereafter; even for a 50% kill-off, the population still recovers (**Figure 8**). While it might be possible to recover faster, doing so causes dangerously large oscillations once the population has returned to its normal levels. This way, the population makes a steady recovery and reaches normal levels while still remaining under control.



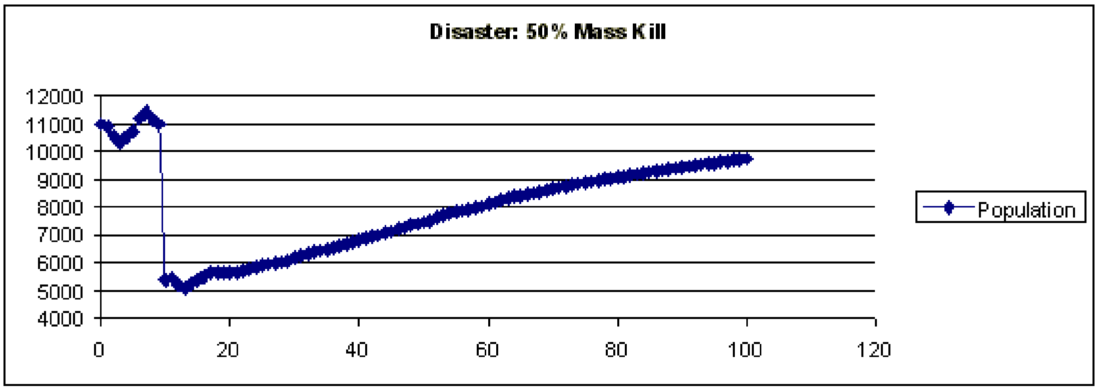


Figure 8. Effects of moderate and severe major disasters (age groups affected equally).

After a natural disaster that kills 10% or even 50% of the very young and very old elephants, the recovery is faster because the young and old are not heavily involved in reproduction (**Figure 9**).

If hunters kill 10% or even 50% of the population over the age of 40, a significant number of reproducing animals are killed, so the recovery is somewhat slower (**Figure 10**).

Our schedule of darting would allow the population to recover from major disasters. Assuming that such disasters occur only rarely, a park using our management policy should have no trouble with population crashes.

Task 5: Justification to the Park Managers

You may well wonder why mathematics is useful in the task of regulating the elephant population in your park. It seems easier to follow a simple set of rules like the following:

- If there are more than 11,000 elephants, dart more than last time.
- If there are fewer than 11,000 elephants, dart fewer than last time.

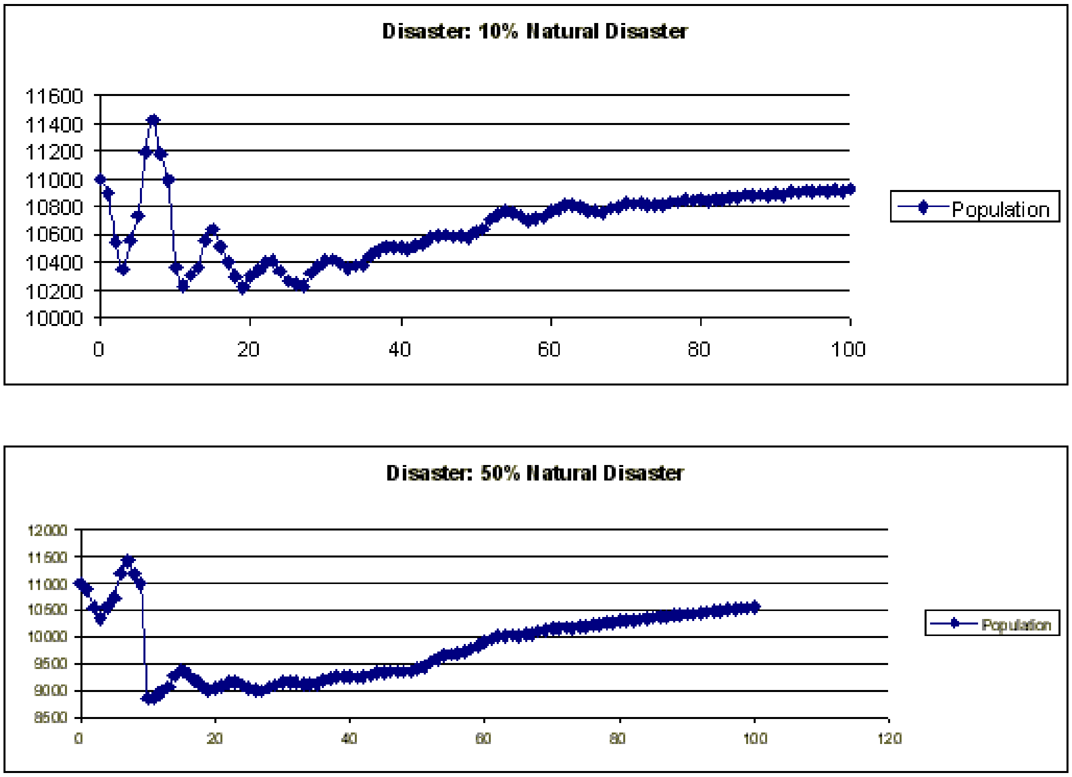
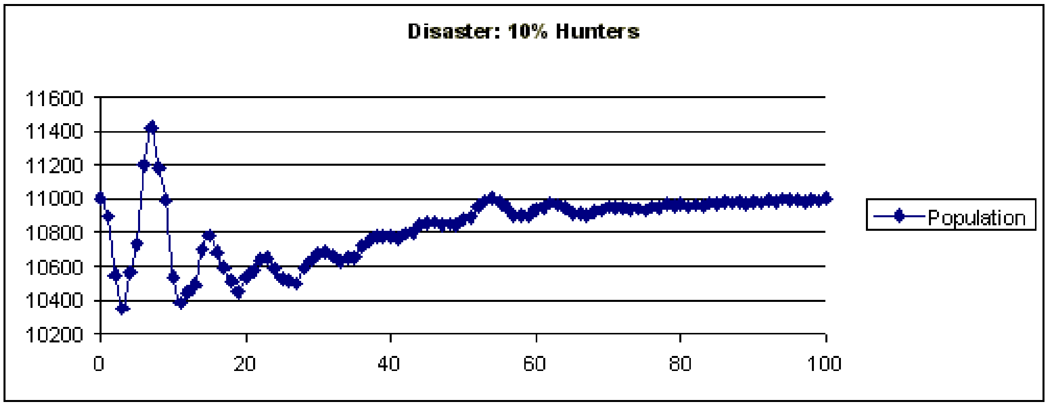


Figure 9. Effects of moderate and severe natural disasters (weakest elephants succumb).

Such a system is simple to understand but difficult to put into practice. For one thing, it is hard to decide on an exact number to increase or decrease the number of darts you are using. The other problem is that changes in the number of dartings does not affect the population for another 22 months. These factors make such a system very problematic in the real world.

Suppose we tried a system of darting a certain percentage of the elephants every two years. If we picked precisely the right percentage, the population would appear to hold steady at 11,000 for a little while, but the fraction of the population that was pregnant would gradually change over time and the population would go out of control faster than the function could compensate. This result can be shown using a simple computer simulation of the population over time.

A better goal than keeping the population constant is keeping the number of elephants born each year constant. Since the rate at which elephants die does not change much, keeping the number of births constant should eventually give a constant number of elephants. Based on elephant birth and death statistics for a healthy herd, we can adjust a healthy population of around 11,000 elephants to a state of equilibrium. By calculating the number of elephants that are less than one year old, we get a good idea of how many elephants were born last year. Dividing by the total number of elephants gives the fraction of the total



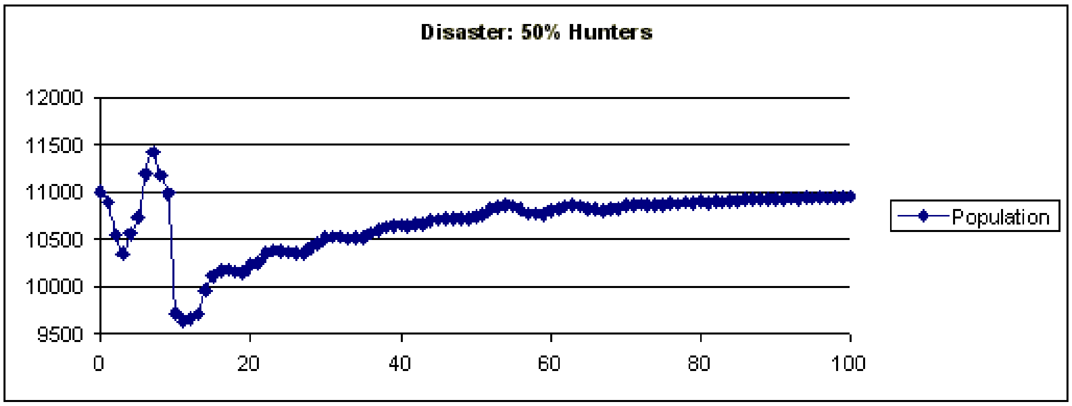


Figure 10. Effects of moderate and severe hunting.

population that must be born each year to keep the population stable.

We constantly have to readjust the number of dartings based on the effect on the future population, for which we provide a formula. We have tried this formula in several simulations and found it extremely adaptable and effective. Its strength lies in the fact that it was derived using sound reasoning; any darting method that does not use mathematics is little better than a wild guess and will not produce satisfactory results. If you use a mathematical model to control your elephant population, you will be satisfied by the long-term behavior of the population. As always, there will be some random fluctuation, but this model provides and effective solution.

Task 6: Generalization

We show that in many cases we can use our model for other parks with different needs.

A key aspect of making our model work is finding an acceptable f_b (the fraction of the population that are newborns) for each target population and set of conditions; this fraction is derived from survivorship data for the individual park's population. Our method forces convergence to the target population.

Suppose that a park has similar conditions but that the death rate among newborn elephants is .35 and the park aims for 25,000 elephants; we find $f_b = 0.046$. This makes sense—the value of f_b must be higher to compensate for the higher death rate, which means that a greater proportion of the population must be newborns in order to maintain the stability of the population.

As a second example, consider a park with a target population of 300 and an infant death rate of 15%. In this case, $f_b = 0.013$—smaller, to compensate for a smaller infant death rate.

Any park with reasonable values for death rates and ideal number of animals should be able to work under this system.

Sensitivity Analysis

For a model incorporating as many parameters as this one does, it is vital to determine which introduce the greatest error. Given a $\pm 10\%$ deviation in the value of the parameter, we calculate the percentage change in the value that the final system converges to. **Table 2** summarizes the parameters that have significant effects; the model is fairly insensitive to the values of other parameters.

Table 2.

Sensitivity of the model to changes in parameters.

Variable	From data	+10%	Equilib. Herd Size	% Diff.	−10%	Equilib. Herd Size	% Diff.
Newborn survival rate	.75	.825	16,200	47%	.675	7,200	−35%
f_b	.0255	.02805	12,100	10%	.02295	9,935	−10%

It is vital to know accurately the newborn survival rate, since the final population is so dependent on this value.

Strengths

- Our methods keep the elephant population under control, which is the main point. The population converges to the ideal number of elephants in a reasonable time.
- Our methods can incorporate various scenarios: contraceptive darting, relocation, compensation for disasters, and application to other similar parks.
- This model is simple enough for the park rangers to understand.
- Our method can produce accurate predictions with very little computer time.

- Our method is robust, so that other variables or situations can be easily introduced.
- After the first five years, under normal conditions the population does not deviate more than 200 elephants from the target value.

Weaknesses

- Our model is somewhat involved, and predictions cannot be generated without a computer.
- The population does not stabilize at exactly 11,000.
- The model responds slowly (though surely) to dramatic changes in the population.
- The method does not allow the relocation of more than 250 elephants per year, which might be possible with a more radical model.

Conclusion

Keeping a dynamic system like an elephant population under control is a very old and difficult problem. It is made more difficult by the long life spans and steady reproductive rates of elephants. We have developed a system that is more humane and more adaptable than simply killing off excess elephants.

References

African Wildlife Foundation. 1998. Wild lives—Elephant. `http://www.awf.org/animals/eleph.html` . Accessed 5 February 2000.

Hanks, John. 1979. *A Struggle for Survival: The Elephant Problem*. Cape Town, South Africa: C. Struik.

Holloway, Marguerite. 1994. On the trail of wild elephants. *Scientific American* 271 (October 1994) (6): 48–50.

Purdy, Judy Bolyard. 1998. Contraceptive Safari. `http://www.ovpr.uga.edu/rcd/researchreporter/spring98/elhant.html` . Accessed 4 February 2000.

EigenElephants: When Is Enough, Enough?

David Marks
Jim Sukha
Anand Thakker
North Carolina School of Science and Mathematics
Durham, NC

Advisors: Dot Doyle and Dan Teague

Statement to the Park Management

We develop a system that uses contraceptive darts as the primary method for elephant population control. This method provides a practical alternative to expensive relocation and unpopular culling. Using a statistical model to simulate the changes in the elephant population from year to year, we determine a darting plan that effectively brings the elephant population down to a stable total population of about 11,000, the park's desired target.

Theoretically, this model should accurately predict the structure and size of the elephant population based on the information provided to us about the elephants, such as birthrates, reproductive activity, and life span. Although we had to determine the elephants' survival rates from a rather small sample of data, the survival rates that we determined matched the general information provided. If more accurate survival rates can be found, the model can be adjusted easily by changing a few parameters.

Additionally, we generalize our model to an adaptive darting method that accounts for random fluctuations due to varying survival rates and birth rates, as well as such external influences as immigration, emigration, and poaching. Thus, despite lack of conclusive data, the darting method will effectively control the population even with the random variations introduced by nature.

This method involves the following basic procedure:

- From a survey of the population, determine the approximate population size, age structure, and survival rates. We estimate these from the sample data provided.

The UMAP Journal 21 (3) (2000) 417–430.

- Feed these data into the mathematical model and from it obtain the initial "dosage" (percentage of females to be darted).
  - If relocation of elephants in not a viable option, base dosage for your park is 57%.
  - If it is possible to relocate about 50 to 300 elephants every year, then only 31% of the reproducing female population needs to be darted.

  The females to be darted can be chosen at random, but measures should be taken not to dart the same female twice nor females who are too old or too young to reproduce, as this would reduce the effective proportion of females treated by the contraceptive. Once darting is complete, it is not necessary to track which individuals have been darted, as darting will not be done again until the current dosage wears off, two years later.

- Every two years, count the population and apply the simple formula given in the technical report. We also provide a separate formula for use if the removal of 50 to 300 elephants per year is anticipated.

Under ideal conditions, the park would continue to use the same initial darting plan. However, the population will naturally experience some deviations from the ideal. When surveys show fluctuations in the population, the provided formulas supply the new proportions needed to correct for the deviations.

Our model also tested the survivability of a population after the elimination of a large proportion of elephants. A large natural disaster or widespread disease might cause such a drop in population. Our tests show that when 80% of the population is killed and when survival rates are reduced by 30% for the next 10 years, there is a statistically significant difference between how quickly the population recovers with and without using contraception. However, the darted elephant population still rebounds if darting is stopped, though with a small lag time.

Concern expressed over the validity of the modeling process, especially when the initial data are not completely accurate, is reasonable given the levels of uncertainty that we are working with. However, no matter what method is used for population control, one must have a relatively good idea of the population structure. Our simulations show that our darting plan is flexible and can accommodate variability or inaccuracies in the initial data. This suggests that our model does not depend as heavily on the initial population structure as other methods of population control might. Of course, the best advice we have for increasing confidence in our model is to collect more data. This would provide the most conclusive evidence for the model's accuracy.

Assumptions

- The number of elephants relocated in the past two years is representative of the actual age structure of the current elephant population. One common practice is to relocate entire family units of elephants at once, which would be generally consistent with this assumption.
- The population in the park never differs greatly from the stable state of the population.
- Elephants mate and give birth at a uniform rate throughout the year.
- The population is sufficiently large that we can compute all relevant quantities concerning the population probabilistically.
- The gestation period can be taken to be two years (as opposed to the given twenty-two months).
- The survival rate within ten-year-wide age groups is roughly uniform.

Analysis of the Problem

The nature of this problem suggests that the population should be modeled by a system of difference equations. The data provided by the park are presented in terms of a discrete age distribution. Since the duration of the darts' effectiveness is given in terms of years rather than a fraction thereof, it is appropriate to approach the problem in terms of a discrete time step, namely $\Delta t = 1$ year. This time step sets iterations at one per year and also stratifies the population into cohorts (age groups) of elephants born in the same year.

Given that all elephants die by the age of 70, the problem is reduced to 70 difference equations, one for each age cohort. Such a system is most naturally represented in terms of a matrix equation

$$P_{n+1} = TP_n,$$

where P_{n+1} and P_n are column vectors with 70 rows, in which the ith element represents the number of elephants of age i. The matrix T is 70×70, and each of the elements in the ith row is a coefficient in the ith difference equation. The matrix representation has a powerful advantage over the system of difference equations: T can be manipulated (e.g., by darting) so that it has an eigenvalue of 1, which corresponds to a stable population and age structure.

A matrix A with eigenvalue λ (a scalar) and eigenvector x has the property that $Ax = \lambda x$. For a general population vector P, as $n \to \infty$, we have $A^n P$ approaches x or some scalar multiple of x. The convergence is especially fast if P is initially somewhat similar to x, although small variations of P from x can cause P to converge to a scalar multiple of x instead of to x itself. This relationship suggests the solution to the dilemma of stabilizing the elephant population: If T is manipulated through darting so that it has an eigenvalue 1, then as it is applied to the population of elephants over time, P will converge to the eigenvector; that is, the population will stabilize.

Determining the Transition Matrix

To determine the elements of the matrix equation, consider the structure of the difference equations. The first reduction in the magnitude of the problem is to consider only female elephants. Given that the sex ratio is "very close to 1:1" for adults as well as for newborns, we can consider only females, knowing that the full population can be determined simply by multiplying by two. Hence, the sum of the elements of the P vector should be close to 5,500. The first element of the P vector is the newborn elephants, age 0. The size of this stratum at iteration $i+1$ depends on only the number of reproducing females. The difference equation for the newborn elephants is then

$$(P_0)_{n+1} = \sum_{i=10}^{60} p_i \cdot (P_i)_n,$$

where p_i is the probability that an elephant in the ith age group has a calf that year and $(P_i)_n$ is the ith element of the nth iteration of P; that is, $(P_i)_n$ is the number of elephants in the ith age group in the nth year of iteration. The value of each of the remaining elements in the P vector is determined only by the number of elephants in the previous stratum that survive into that year. This can be written as

$$(P_i)_{n+1} = s_{i-1} \cdot (P_{i-1})_n,$$

where s_i is the probability that an elephant of age i will survive until the next year. This suggests that T is of the form

$$T = \begin{pmatrix} 0 & p_1 & p_2 & 0 \\ s_0 & 0 & 0 & 0 \\ 0 & s_1 & 0 & 0 \\ 0 & 0 & s_2 & 0 \end{pmatrix},$$

but much larger (70×70). Further simplification is desirable.

If P is "close" to an eigenvector, then with each iteration the same numbers of elephants grow into the next age level as they did the previous year; that is, if P is nearly stable, then the population structure should remain relatively constant from year to year. This also means that a larger stratum, say 10 years, has a predictable age distribution. Namely, if a stratum has c elephants growing into it every year with a constant survival rate s over the stratum, then the total number of elephants in the interval is

$$N = c\left(1 + s^1 + s^2 + \cdots + s^n\right),$$

where n is the width of stratum N. The proportion of elephants growing out of stratum into the next stratum is given by

$$\text{Growth} = \frac{cs^n}{c(1+s^1+s^2+\cdots+s^n)} = \frac{s^n(1-s)}{1-s^{n+1}}.$$

Thus, in the steady state, several years of elephants can be grouped together without any loss of information. For the purposes of further discussion, we assume (and verify later) that P is indeed sufficiently close to the eigenvector, and thus we collapse the elephant population into 8 strata, the newborns plus one for each decade up to age 70. The T matrix, now only 8×8, is of a slightly different form, namely,

$$\begin{pmatrix}
0 & 0 & p_2 & p_3 & p_4 & p_5 & p_6 & 0 \\
s_0 & s_1(1-g_1) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & s_1 g_1 & s_2(1-g_2) & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & s_2 g_2 & s_3(1-g_3) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & s_3 g_3 & s_4(1-g_4) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & s_4 g_4 & s_5(1-g_5) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & s_5 g_5 & s_6(1-g_6) & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & s_6 g_6 & s_7
\end{pmatrix},$$

where $g_i = \left[s_i^n(1-s_i)\right] / \left[1-s_i^{n+1}\right]$ is the proportion of elephants that move out of stratum i and into stratum $i+1$ each year. This leaves the determination of survival rates s, probabilities of birth p, and the initial population matrix P to be determined.

Current Age Structure

To determine the survival rate as a function of age, we look to the data provided by the park. From the past two years, we have the sex and approximate ages of the elephants transported out. Under the assumption that elephants were removed fairly uniformly, we take these data to be an accurate representation of the park's overall elephant population. We extrapolate from these data the age distribution of the elephants in the park.

We assume that the elephant population is reasonably stable, in particular, that the overall age distribution of the population is the same for both years. Additionally, since the sex ratio is "very close" to 1:1, we can treat a distribution of one sex as representative of the population's age distribution. We have four samples, one from each sex for each year; we combine these four samples to obtain the relative frequency of elephants at each age (which we scale such that the total number of elephants is 11,000, the park's total population). This distribution is shown in **Figure 1**.

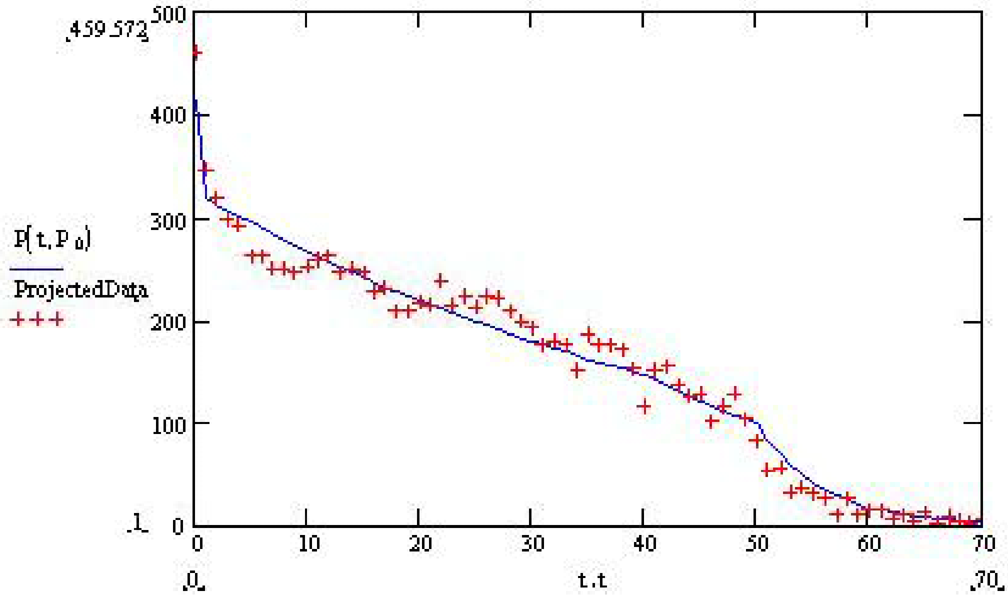


Figure 1. Projected age structure.

Survival Rate

In addition to the age distribution, we are also interested in determining the survival rate of elephants in a given cohort. We begin by following a cohort over time, plotting the number of survivors each year. For a fairly stable population, this relationship is identical to the age distribution. That is, if a population is stable, the age structure is not changing significantly; so as a cohort ages, its size must change to fit the population's age structure.

Since the age structure and cohort survivor data are nearly the same, we use the previously determined age structure to determine the elephants' survival rates as well. Survival rate is defined as the probability that an elephant at a given age survives to the next year. For example, if an elephant has a survival rate of 75% at age 0, the probability that it survives for the next year is 0.75. For large mammals such as elephants, we expect the survival rate in the middle of an elephant's life cycle to remain relatively independent of age while being much lower in very young and very old elephants. Since survival rate governs the change in population from one year to the next and is proportional to the current population, we expect the age structure data to be exponential. The graph in **Figure 1** verifies this prediction.

We plot the natural log of the number of elephants versus age (**Figure 2**) and fit lines through the four major sections of the data, from which we determine the survival rates for four major sections of the population. (**Table 1**).

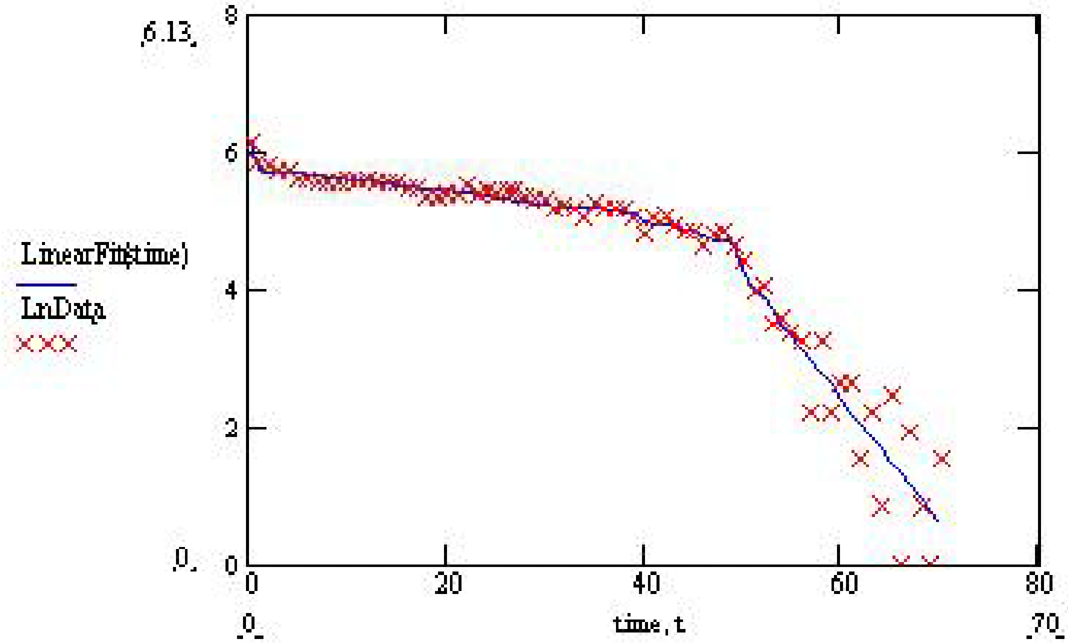


Figure 2. Natural log of elephant data vs. age.

Table 1.
Elephant survival rates.

Age Group (years)	Survival Rate (% / year)
0–1	75
1–50	98
51–60	96
61–70	82

Probabilities of Birth

The information provided by the park suggests that the reproductive rate is constant over all reproducing age groups except the teenage group, where not all of the elephants are reproducing. On average, a female cow produces a calf every 3.5 years with twins occurring 1.35% of the time, indicating that the probability of any given female producing a female calf in a given year is $0.5(1.0135)/3.5 \approx 0.145$. For the teenage group, we assume that about one-third begin conceiving when they are 10, another third when they are 11, and the remaining when they are 12. However, since it takes about two years from conception to birth, the elephants do not actual have young until they are 12. Taking all this into account, the p value for the teenage stratum is 0.7 times that of the other groups. This completes the matrix T (before we start to consider darting).

Introducing the Contraceptive Dart

If the park management darts female elephants at random, then we can assume that the same proportions of each reproducing stratum are sterilized for the two-year period. Hence, if the management keeps some proportion $1-q$ of the population sterile at any given year, then the transition matrix is the same as T above except that the first row has a factor of q in front of the probabilities for birth. The parameter q is the proportion of females reproducing, and this is the value that can be altered to allow T to have an eigenvalue of 1. For an eigenvalue of 1, T must have the property that the determinant $|T - I|$ (an eighth-degree polynomial in q) is 0, where I is the identity matrix. Once q is known and the appropriate T matrix is constructed, an eigenvector can be found quickly. This allows for speculation as to the desirable steady-state elephant population.

The Ideal Solution

Using the values for s and p determined above, the proportion q of females that should be kept reproducing to keep the population stable is about 43%, that is, 57% of the females should be on contraceptives. The appropriate eigenvector associated with a population of 5,500, as well as the extrapolated population estimated from the given data, are shown in **Table 2**.

Table 2.

Eigenvector and extrapolated population.

Cohort	0	1–10	11–20	21–30	31–40	41–50	51–60	61–70
Eigenvector	207	1422	1162	950	776	581	221	181
Extrapolated population	222	1396	1185	1077	833	615	144	27

A measure of the difference of the estimated population from the eigenspace is the cosine of the angle between them. Using the dot product $u \cdot v$, the cosine of the angle between the two vectors is

$$\cos\theta = \frac{EV \cdot EP}{\sqrt{EV \cdot EV}\sqrt{EP \cdot EP}}.$$

For the population estimated from the initial population, we get $\theta = 5.3°$, so the initial population is indeed already close to the eigenvector and our approximation by using 8 strata instead of 70 is valid.

Having set q so that T has an eigenvalue of 1, the matrix is left-multiplied on the population vector. It takes two years for the contraceptive plan to begin to work, due to the two-year gestation period. Using the extrapolated population vector from **Table 2** and assuming the two-year lag period has already passed, **Figure 3** shows the model's prediction of the population over 60 years.

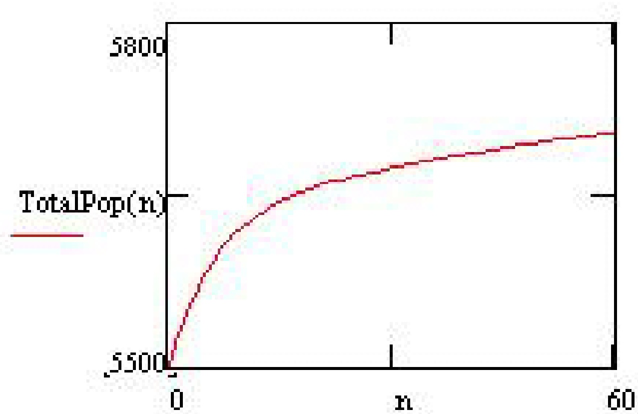


Figure 3. Elephant population over a 60-year period.

The contraceptive rate planned causes a convergence of a population, but it takes a long time. Additionally, the population converges to a higher multiple of the desired eigenvector, with the sum all the female elephants at about 5,700.

The solution to this dilemma is to use not just a single darting plan. Rather, if a more aggressive contraceptive program is used for the first 15 years, the eigenvalue of the matrix will decrease. If the eigenvalue is less than 1, the population eventually begins to drop as it converges to an eigenvector that gets smaller with each iteration (although initially there may be an increase). Once the population has dropped sufficiently, switching to the original darting plan causes convergence to a population distribution with the total number of elephants closer to the desired number (**Figure 4**).

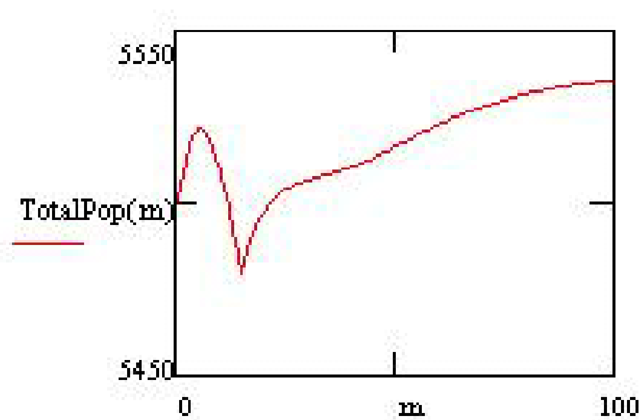


Figure 4. Elephant population over a 60 years with two-phase contraceptive plan.

One particular solution is initially to inoculate about 60% of the population instead of 57%. This small change in the initial 15 years leads to a less variable convergence, where the total elephant population is never more than 100 elephants from the desired 11,000.

The Adaptive Solution

The complication with the two-phase solution is that the exact values for the inoculation are very dependent on properties of the matrix itself, namely, survival rates and birth rate probabilities. Also, more substantial random perturbations in the population, as caused by such phenomena as immigration, emigration, and poaching, can make a static inoculation plan ineffective. Thus, it makes sense to develop a plan that depends on how the population is reacting. If the park can determine the amount of darting needed in the ideal case, then using that value as a base, the park can adjust the actual number of inoculations as required by year-to-year changes in the population.

Given a population at year n, the park would like to know what proportion of females to inoculate. However, changing the proportion of elephants inoculated in year n does not have an effect on the birth rate until year $n+2$, because of the two-year gestation period. In addition, the exact proportion inoculated cannot be adjusted every year independently of the previous year, as the effect of the dart lasts two years.

Thus, there are two possible plans:

- Dart every year, allowing management to raise the levels whenever necessary. If the contraceptive rate every needs to be substantially lower, there will be a one-year lag before a new darting regimen has an effect.
- Dart only every other year, meaning that the management refrains from raising the levels during the off years; but if the dosage needs to go down, there is probability one-half that it can occur immediately.

Both plans have advantages, but the second plan uses substantially fewer darts and will in general be cheaper and require less work to implement.

As the female population changes from 5,500, either due to an eigenvalue not being 1 or to natural perturbations, the number under contraception must be adjusted. The desired proportion $p(N)$ sterile as a function of population N should have the properties that when $N = N_0 = 5{,}500$, $p(N_0) = p_0$ (the value necessary to give the projected T matrix an eigenvalue of 1), and $p(N)$ decreases as N decreases and increases as N increases. To allow for ease of generalization, the amount that p varies should depend on the percentage difference between N and N_0: The effect should be small for small differences in N but should grow quickly enough to constrain N if changes in N are too great. A linear relationship grows too rapidly, suggesting a natural logarithm function of the form

$$p(N) = p_0 + c\,\mathrm{sgn}\left(\frac{N - N_0}{N_0}\right) \cdot \ln\left(\left|\frac{N - N_0}{N_0}\right| + 1\right),$$

where c is a constant that determines how reactive the darting is to changes in the population. We find that values of c ranging from 3 to 5 work well in keeping the population stable (see the **Appendix**). We also find from our simulations that constraining p between some upper and lower bounds increases the stability of the population; a reasonable constraint for p is $0.3 < p < 0.7$.

Contraception with Relocation

If the park managers have the option of removing some elephants in addition to darting, the nature of the problem changes a bit. Consider a population where some number m elephants are to be removed. If they are taken equally from the various strata, then the transition matrix T should be constructed so that it has an eigenvalue $1 + m/N$. This again will allow for a steady state, as each year there will be m more elephants but that many will be relocated. If m is constant, then the problem is solved, as the solution is exactly the same as before, with a slightly different value for p—31%, which is much lower than that amount of contraception use otherwise.

A Disaster

A warranted concern regarding darting is what happens immediately following a natural disaster. We examine the effects by evolving a population for 25 years and then simulating a large natural disaster. This disaster kills 80% of the population and is then followed by a 30% reduction in the survival rates for 10 years. **Figure 5** shows the rebound of a population that begins to reproduce immediately, while **Figure 6** is for a population that must first go through a lag period due to the contraceptives. While the random effects cause some variations depending on the simulation, the overall trend is that the population without contraceptives bounces back faster. Based on 10 simulation runs, the mean population at year 120 with no contraceptive use is 1,246 (SD = 331), while it is only 1,009 (SD = 286) for a group on contraceptives.

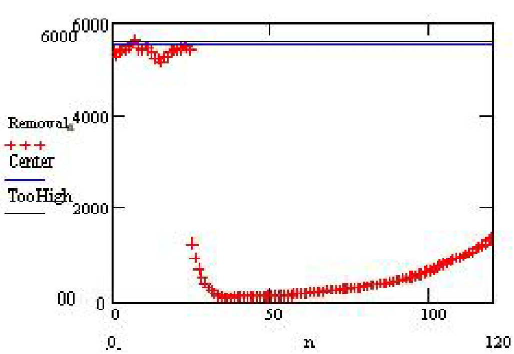


Figure 5. Effect of a disaster on an undarted population.

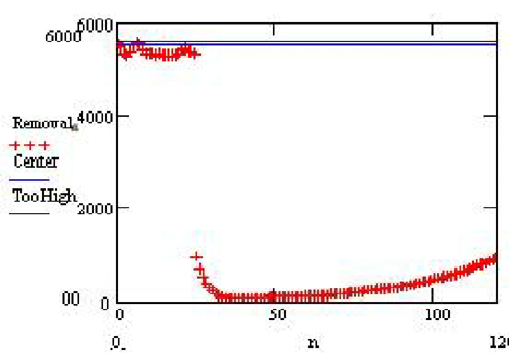


Figure 6. Effect of a disaster on a darted population.

A one-sided t-test of the difference, with df = 17.62, gives a P-value of 0.052, on the border of significance at the 5% level. The opponents of darting may be correct in concerns about an impeded ability of the elephants to grow back. However, the elephant population will still return, if at a slightly retarded rate. Controlling the population without culling elephants seems to justify the risk.

Generalizations

The key to stabilizing a different population at another park is to find the survival rates and birth probabilities, then determine the right value of p that allows T to have an eigenvalue of 1. This process is entirely independent of the actual size of the population, depending on only its age distribution. For small populations, the approximation used to simplify the matrix from 70×70 to 8×8 may begin to break down, but that can be fixed by simply expanding the matrix, which requires just more computer time.

Strengths and Weaknesses

Weaknesses

- Our model for survival rates and age structure depends heavily on the elephant removal data. If these data are not representative of the overall age distribution, then the final population that the model predicts may deviate slightly from the actual value. A more meaningful conclusion on the current age structure cannot be obtained without additional data.
- Our transition matrix considers elephants in 10-year age groups rather than 1-year cohorts. This simplification greatly reduces the size of the transition matrix and allows for quicker calculations, but the approximation may introduce slight inaccuracies. The inaccuracies grow if the population distribution is drastically different from the ideal distribution.
- Elephant populations are discrete quantities, but we approximate them with continuous values. For extremely small populations, this approximation may no longer be valid, especially in the older age groups where there are already very few elephants.
- Our initial model, without adjusting the level of contraceptive darting, is sensitive to changes in survival rates; different values for those can cause the population to converge to a different final value. The modified model that makes adjustments to the level of darting is more capable of handling slight changes in survival rate, but a significant change can still alter the final results.

Strengths

- The final model handles small random fluctuations in the population quite well. These fluctuations add a reality check because they reflect possible error in the park managers' estimate of the population size. The population remains within a reasonable interval around the ideal population, which means the model is not very sensitive to variations in population size.

- The model considers the possibility of some elephants being relocated each year. Relocation when feasible is a preferred method of population control, but the model is not dependent on this possibility.
- Our model can be modified easily to accommodate other parks with different populations and survival rates.

References

World Wide Fund for Nature. 1997. Conserving Africa's elephants—Conservation inside protected areas. `http://www.panda.org/resources/publications/species/elephant/elephant3.html` . Accessed Feb. 5, 2000.

Appendix: Simulation of Adaptive Darting

We determine useful values for c. After each iteration of the simulation, each element in P is multiplied by a random number $\epsilon \sim N(1, 0.01)$. This introduces an effect on the order of 1–2% variation from the predicted value. These effects can be due to errors in the matrix, elephant movements, poaching, or other random effects. We find that effective values for c range from 3 to 5, giving the parks much flexibility in estimating the how much the female population deviates from the desired 5,500.

However, due to the lag involved in the contraceptive's effects, the population meanwhile may deviate farther from 5,500. Placing upper and lower limits on the contraceptive dosage minimizes this problem. Just as with the values for c, the simulation does not change too greatly with different values for the upper and lower bounds on the contraceptive dosage p; a reasonable range seems to be $0.3 < p < 0.7$. The graphs shown in **Figure A1** are eight consecutive random trials, the first four with constraints, the next without. With the exception of the last simulation with constraints, the constraints restrain variability a little but not a lot, suggesting that the park need not worry about exact calculations. Randomness does cause the convergence of the model to disappear. The advantages are that this plan can be started immediately and does not rely on perfectly uniform natural conditions.

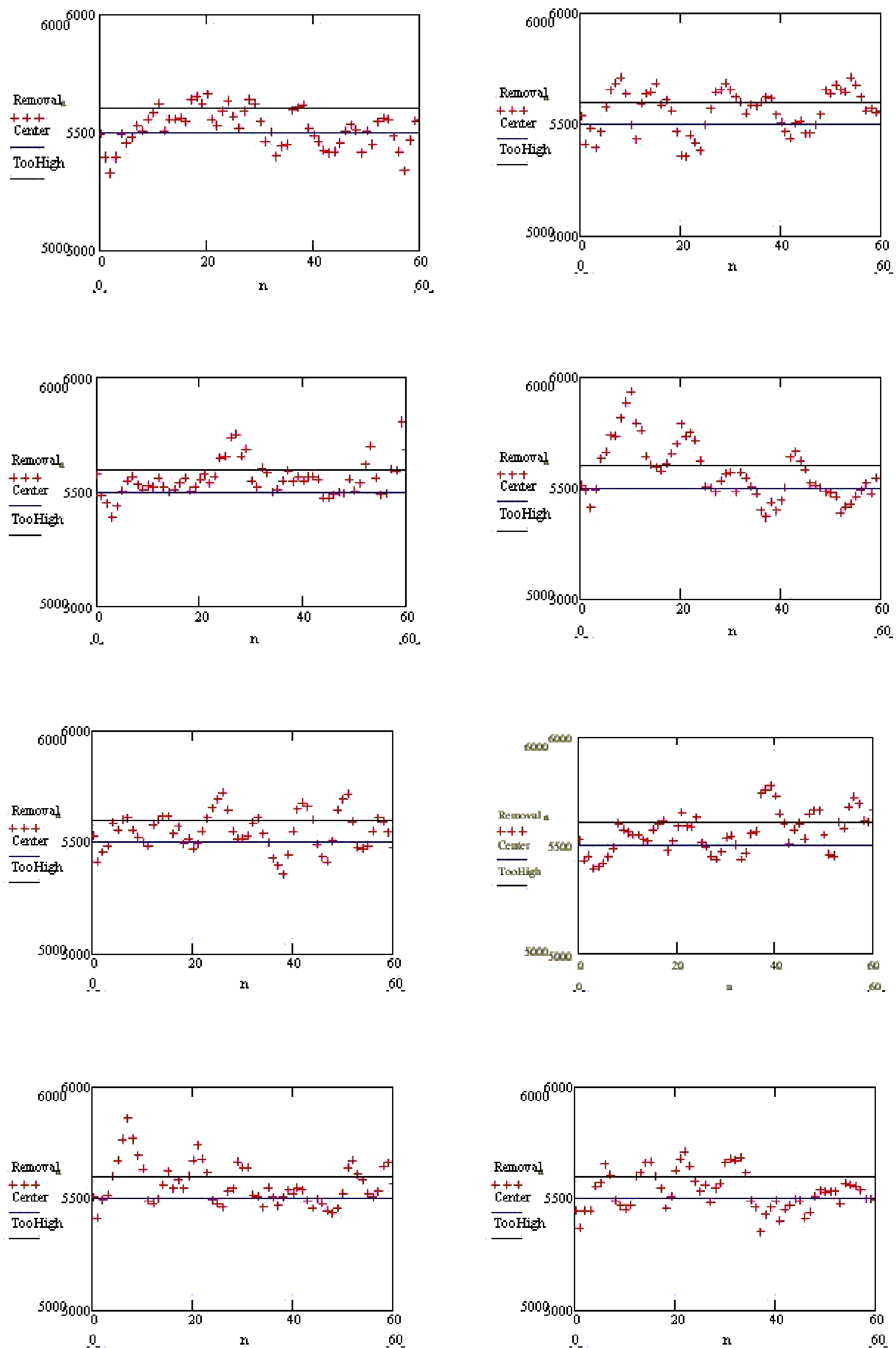


Figure A1. Adaptive contraceptive use with random fluctuations.

Judge's Commentary: The Outstanding Elephant Population Papers

Gary Krahn
Department of Mathematics
U.S. Military Academy
West Point, NY
ag2609@exmail.usma.army.mil

Introduction

The judges were very impressed by the breadth of insight revealed by the "modelers" in this the second year of the ICM Contest. Each of the six required tasks in the problem were individually weighted; however, papers were ultimately evaluated on their overall effectiveness to formulate a policy that would solve the overpopulation problem and create a healthy environment for a herd of 11,000 elephants.

The Problem

At first glance, the information and data provided in the problem statement appear to be sufficient to construct a model to capture the population growth of the elephants under specific control measures. This problem, however, was not clear-cut. As the contestants formulated and refined their assumptions, they confronted the complexities typically associated with an open-ended problem. The initial task was to develop and use a model to investigate how the contraceptive dart might be used for population control. The modelers, however, do not know the initial age structure. Nor are they privy to the implementation procedures of culling, relocation, darting, and the cultural issues surrounding population control.

The UMAP Journal 21 (3) (2000) 431–434.

The Science

Papers were evaluated based on an understanding and application of science, model development, and analysis. Knowledge of the elephant's life-cycle and survival characteristics was important in the design of an appropriate solution to this problem. Environmental science provided the modelers with the foundation to make "realistic" and appropriate assumptions. More important, it gave teams confidence in the data provided and in the results of their models. In essence, understanding the science of the biological and environmental data transformed this problem into a real-world application. Many teams, through research, verified that after 5 years of age, elephants live in relative safety—there is a low rate of terminal diseases, accidents are rare, and there are few natural predators. In addition, they discovered that deaths of elephants often occurs after 60 years of age because of eating complications. Teams that gained an understanding of biological issues affecting elephants also achieved better "control" of the information in the problem statement. As a result, the top interdisciplinary teams were able to find insights into the significant parameters that influenced the elephant population growth.

The Model

Some teams constructed an analytic model, some used population models found in the literature, and others developed simulations to replicate the real world behavior. All teams used some simplifying assumptions to reduce the scope of the problem. The judges thought it was important to keep the assumptions reasonable and to avoid making unnecessary assumptions. Several teams examined (or constructed) different population models simultaneously to verify their work and to gain perspectives on how to adapt established modeling techniques for this particular problem. A few teams effectively simplified the problem by modeling only the female population. More than a few teams used several solution techniques (the Leslie matrix and a simulation) and compared the results. Other teams constructed models that captured the dynamics of individual elephants and compared that to models that grouped elephants into categories. The judges were heartened by the number of teams that attempted to validate the models.

In Task 4, the modelers were asked to investigate the affect of disease and uncontrolled poaching after darting. Teams that used several modeling approaches almost always discovered that the population would oscillate for several years after a dramatic population change and then would recover. Simulations were effectively used to reveal this phenomenon and graphical techniques were able to display this result very clearly.

The Analysis

The mathematics required to explore the dynamics of the population growth did not require sophisticated methods. Difference equations, discrete dynamical systems, differential equations, transformation matrices (Leslie model), and computer simulations were applied with great success by many of the 69 teams in the contest. Some teams examined the long-term behavior of the system using the eigenvalues of the transition matrix—a very nice application of matrix algebra. The judges were delighted that the top teams discovered that their results were insensitive to assumptions about the initial age structure. Computation and calculation were not the most important features for successfully solving this problem. The reasoning process, modeling, and problem solving were of much greater importance.

Interdisciplinary

Again, the characteristic of a strong paper was the knowledge of environment science and resource management and application of valid modeling concepts, along with terminology that explained the analysis, outcomes, and recommendations. The top papers not only conducted a thorough analysis but also shared their method of reasoning in sufficient detail. The problem statement revealed that park officials were very skeptical about mathematical modeling. Therefore, it was essential to outline the modeling procedures and the implications of any assumptions. The analyst's credibility (essential for the eventual implementation of the model) could be enhanced by revealing knowledge about the elephant life-cycle, discussion of the advantages and disadvantages of models/simulations, and sharing with the park officials an appreciation of the complexity of the problem. Some National Parks in South Africa cover over 2 million acres, larger than the state of New Jersey (and no turnpike!). Counting elephants in these rugged areas is not simple—it necessitates historical evidence and statistical inference. Determining the age and sex distribution of elephants can be extremely difficult especially during periods when external forces are changing the natural equilibrium of the herds. Discussing these considerations of the ecosystem with the Park Officials was an important element of Task 5—increasing the confidence of park managers. A team's lack of appreciation for the complex environment was often revealed in the manner in which the darting would be implemented. Some teams suggested counting elephants every year and only darting elephant of specific age groups—probably an impossible undertaking. Other teams realized that it would be impossible to "tag" darted elephants and suggested darting every two years to help eliminate the problem of darting the same elephant several times in a given year.

The interplay between darting and relocation was the theme of Task 3. It was interesting how some teams believed that relocation could be done very easily and other teams wrestled with the cost and complications of relocating hundreds of elephants—hundreds of big, heavy, cumbersome, stubborn elephants. The understanding of resource management was a critical ingredient in solving this problem. The top teams did this very well.

Presentation

Clarity of presentation is essential to good research and analysis and it provides the ability to effectively influence the decision making process. Many teams this year presented very clear and concise support of their work. The stronger teams carefully created the appropriate mixture of words, graphs, algorithms, and analysis to present their reasoning and recommendations.

Often, the results for a specific task were spread throughout the paper and were not confined to a particular section of the paper. Over the years, however, modeling teams have continued to place a greater emphasis on the write-up. This has been a very pleasant trend to witness.

Conclusion

Almost every team felt comfortable transitioning their work to other possible scenarios. They revealed a confidence in generalizing their analysis and adapting it to specific situations.

This problem was successful because to write a top paper required an understanding of science, research, and mathematics. The best teams revealed the value of solving a problem from an interdisciplinary perspective. The top three papers are remarkable efforts to solve an open-ended problem in a very shot period of time! Congratulations to all the interdisciplinary teams and especially the three "outstanding" teams.

About the Author

Gary Krahn received his Ph.D. in Applied Mathematics at the Naval Postgraduate School. He is currently the Deputy Head of the Dept. of Mathematical Sciences at the U.S. Military Academy at West Point. His current interests are in the study of generalized de Bruijn sequences for communication and coding applications. He enjoys his role as a judge and associate director of the ICM.

About the Problem Author

Chris Arney
Gary Krahn
Department of Mathematics
U.S. Military Academy
West Point, NY
`ad6819@exmail.usma.army.mil`
`ag2609@exmail.usma.army.mil`

In the second year of ICM, the contest directors wanted a problem involving resource management as it relates to modeling in environmental sciences. In addition, we wanted the problem to involve data analysis, to be realistic, and to be open-ended where no solution is readily available. Our search converged to the elephant problem of Professor Tony Starfield—and we believe our search was successful.

Professor Anthony M. Starfield in the Dept. of Ecology, Evolution, and Behavior, University of Minnesota, carefully guided us in composing this problem. He is an applied mathematician who enjoys using mathematics and computers to help solve "real-life" problems. Prof. Starfield has his Ph.D. from the University of the Witwatersrand, Johannesburg, and has worked on many problems in wildlife conservation, in particular, modeling of populations and ecosystems. For 20 years, he has built models to aid management decisions in the game parks of Southern Africa. His current research work is in two separate areas.

- He investigates how decisions are made in conservation biology and attempts to develop models that feed into a formal multi-objective decision process that reflects both the uncertainty inherent in conservation problems and the various interests of the players in these decisions.
- He develops new paradigms for modeling ecosystem dynamics. This approach has been applied to forest succession in Minnesota, elephant-tree dynamics in Zimbabwe, and global warming on Alaskan tundra.

Tony Starfield is also the co-author of *How to Model It: Problem-Solving for the Computer Age* and *Building Models for Conservation and Wildlife Management*.

The UMAP Journal 21 (3) (2000) 435–437.

Our question combined two components that Prof. Starfield has incorporated into his work:

- using mathematical models effectively in ecology and conservation biology, and
- using the creative aspect of modeling as a logical and practical process.

If you compare the ICM problem with the problem that Tony Starfield has been investigating (see below), you will see that only the names have been changed to protect the innocent. As we coordinated with Tony, we found that he spends considerable time in South Africa solving important problems. Here is the problem that Tony suggested:

> The Kruger National Park in South Africa has tried to maintain a steady elephant population. Their policy is to keep the number of elephants fixed; and for the past 20 or more years they have attempted to count the total population each year, then remove whole herds to keep the population stable. This operation has involved shooting (for the most part) and occasionally relocating elephants every year. There has been a public outcry against the shooting of elephants, and it is not feasible to relocate large number of elephants. A contraceptive dart has been developed that will prevent a mature elephant cow from conceiving for a period of two years. How can darting help control the population of elephants?

Tony suggested that that overall task is to develop and use models to investigate how the contraceptive dart might be used for population control. He then crafted a series of tasks to help guide the students.

We thought it would be best not to specify the Kruger Park, to allow for greater generality. Because of the assumptions that have to be made, this is an open-ended problem. This problem was exciting because it was accessible to students with a variety of backgrounds, without losing the "real-world" application. It also required teams to do research to model the situation appropriately. A goal was to provide an opportunity to have students discover that modeling skills can allow them to make significant contributions to society. We hope that a little of Tony Starfield's commitment to our environment and modeling rubbed off on those involved with this problem.

References

Starfield, A.M., K.A. Smith, and A.L. Bleloch. 1990. *How to Model It: Problem-Solving for the Computer Age.* New York: McGraw-Hill.

Starfield, A.M., and A.L. Bleloch. 1991. *Building Models for Conservation and Wildlife Management.* 2nd ed. Edina, MN: Burgess Press.

About the Author

Anthony Starfield received his Ph.D. at the University of Witwatersrand, Johannesburg. He is an applied mathematician who enjoys solving environmental problems. For 20 years, Tony has been working with engineers to build models to aid management decisions in the game parks of Southern Africa. What was essentially a hobby grew into a career. Today he would be described as an ecological modeler. Currently, he is a Professor in the Department of Ecology, Evolution, and Behavior at the University of Minnesota College of Biological Science. His home page is `http://biosci.cbs.umn.edu/eeb/faculty/StarfieldAnthony.html`.

United States Postal Service

Statement of Ownership, Management, and Circulation

1. Publication Title	2. Publication Number	3. Filing Date
The UMAP Journal	0197-3622	09/16/00

4. Issue Frequency	5. Number of Issues Published Annually	6. Annual Subscription Price
Quarterly	4 + annual collection	$64.00

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*

57 Bedford St., Ste. 210, Lexington, MA 02420

Contact Person: Kevin Darcy

Telephone: 781-862-7878

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

same

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

Solomon Garfunkel, 57 Bedford St., Ste.210, Lexington, MA 02420

Editor *(Name and complete mailing address)*

Paul J. Campbell, Beloit College, 700 College St., Beloit, WI 53511

Managing Editor *(Name and complete mailing address)*

Pauline Wright, 57 Bedford St., Ste. 210, Lexington, MA 02420

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
Consortium for Mathematics	57 Bedford St., Ste. 210
and Its Applications, Inc.	Lexington, MA 02420
(COMAP Inc.)	

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address
N/A	

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*

The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:

☒ Has Not Changed During Preceding 12 Months

☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

PS Form **3526**, September 1998 *(See Instructions on Reverse)*

13. Publication Title The UMAP Journal			14. Issue Date for Circulation Data Below 9/29/00	
15.		Extent and Nature of Circulation	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*			1275	1820
b. Paid and/or Requested Circulation	(1)	Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)*	N/A	N/A
	(2)	Paid In-County Subscriptions *(Include advertiser's proof and exchange copies)*	N/A	N/A
	(3)	Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	80	75
	(4)	Other Classes Mailed Through the USPS	N/A	N/A
c. Total Paid and/or Requested Circulation *[Sum of 15b. (1), (2),(3),and (4)]*		▶	80	75
d. Free Distribution by Mail *(Samples, compliment ary, and other free)*	(1)	Outside-County as Stated on Form 3541		
	(2)	In-County as Stated on Form 3541		
	(3)	Other Classes Mailed Through the USPS	80	95
e. Free Distribution Outside the Mail *(Carriers or other means)*			-0-	-0-
f. Total Free Distribution *(Sum of 15d. and 15e.)*		▶	80	95
g. Total Distribution *(Sum of 15c. and 15f)*		▶	160	170
h. Copies not Distributed			125	590
i. Total *(Sum of 15g. and h.)*		▶	285	760
j. Percent Paid and/or Requested Circulation *(15c. divided by 15g. times 100)*			50%	44%

16. Publication of Statement of Ownership
☒ Publication required. Will be printed in the third issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner: Solomon Garfunkel — Date: 9-19-00

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Periodocals authorization.