

The Flux Measure of Influence in Engineering Networks

by

Kyle Michael Schwing

B.S., Mechanical Engineering (2008)
B.S., Global Economics and Modern Languages (2008)
Georgia Institute of Technology

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

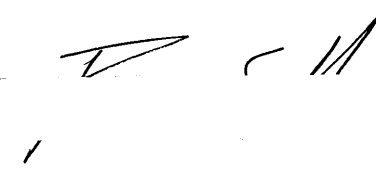
September 2009

© Massachusetts Institute of Technology 2009. All rights reserved.

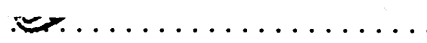
Author

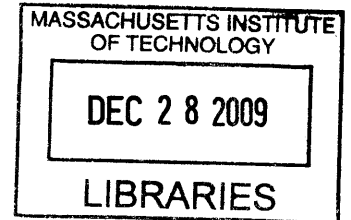
Department of Mechanical Engineering
August 7, 2009

Certified by


Franz S. Hover
Assistant Professor
Thesis Supervisor

Accepted by


David E. Hardt
Graduate Officer, Department of Mechanical Engineering



ARCHIVES

The Flux Measure of Influence in Engineering Networks

by

Kyle Michael Schwing

Submitted to the Department of Mechanical Engineering
on August 7, 2009, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

The objective of this project is to characterize the influence of individual nodes in complex networks. The flux metric developed here achieves this goal by considering the difference between the weighted outdegree and indegree of a node. This technique differentiates among nodes that traditional centrality measures treat as identical units. The behavior and proper interpretation of the flux metric are demonstrated on a variety of weighted and directed networks. Simulations of fluid flow, opinion sharing, epidemic dynamics, and resource allocation reveal the practical capabilities of the flux metric. An engineering design challenge may also be framed as a network analysis problem so that the flux metric contributes to understanding the relationships among the system's subcomponents and objectives. A case study that investigates the design of autonomous underwater vehicles (AUVs) for use in the offshore oil and gas industry demonstrates these insights. In all of the applications explored here, the flux metric consistently emerges as a reliable indicator of the influence of a node.

Thesis Supervisor: Franz S. Hover
Title: Assistant Professor

Acknowledgments

I would like to thank my advisor, Professor Franz Hover, for his exceptional support and patience. His insight, advice, and encouragement made my time at MIT extraordinary. I would also like to recognize Professor Pierre Lermusiaux for offering his experience in complex systems and providing an extra opinion whenever necessary. Finally, the daily logistics of Hover Group would collapse without the efforts of Sheila McNary.

I will forever be indebted to Professor G. Paul Neitzel, my fluid mechanics professor at the Georgia Institute of Technology and coauthor of “Control-volume momentum flux: getting the sign correct.” His willingness to consider the offhand suggestion of a sophomore caused a great flux in my academic career that lasts to this day. Two of my graduate professors, Professor Radhika Nagpal of Harvard University and Professor Hank Marcus of MIT, very kindly permitted me to develop aspects of this thesis as part of my classwork and taught me that networks link infinitesimal proteins, mighty ships, and everything in between.

I would like to recognize all of the participants in the survey on the design of an AUV for the offshore oil and gas industry. Their professional expertise and opinions were invaluable to me and greatly enriched this project.

Matt Greytak, Josh Taylor, Brendan Englot, Charlie Ambler, Lynn Sarcione, Kyle Schmitt, and Eric Gilbertson have been much more than lab mates. Their intelligence, kindness, and humor have contributed to this work in a thousand intangible ways.

I would like to thank the benefactors of my Pappalardo Fellowship. This generous grant gave me the liberty to adapt my research to my own interests and schedule. These opportunities were truly invaluable, and had a significant positive impact on my MIT experience.

There is absolutely no way that I would have reached MIT, much less left with a degree, without the support of my family. Thank you for encouraging me through the worst days, celebrating after the best days, and loving me every day.

Contents

1	Introduction	15
1.1	Motivation: Dynamic Phenomena on Networks	15
1.1.1	Mathematical Graphs	16
1.2	Prior Relevant Work in Network Analysis	17
1.2.1	Degree	19
1.2.2	Betweenness	20
1.2.3	Eigenvector Centrality	22
1.3	Problem Statement: Identifying Influential Nodes and Their Role in Engineering Design	24
2	Introducing the Flux Metric	25
2.1	Concept of Flux	26
2.2	Comparison of Flux to Previous Metrics	27
2.2.1	Flux and the Secretary Paradox	27
2.2.2	Network Models	29
2.2.3	Unique Information Revealed by the Flux Metric	29
2.3	Exploring the Flux Metric	35
2.3.1	Comparing Flux Values	35
2.3.2	Flux in Graphs of Varying Geometry	37
2.3.3	Flux Neighborhoods	41
3	Applications of the Flux Metric	45
3.1	Synchronization	46

3.1.1	Relevant Previous Work	47
3.1.2	Model Description	48
3.1.3	Engineering Analysis	50
3.2	Infection	53
3.2.1	Relevant Previous Work	54
3.2.2	Model Description	56
3.2.3	Engineering Analysis	57
3.3	Resource Sharing	61
3.3.1	Model Description	62
3.3.2	Engineering Analysis	63
4	Engineering Design through Network Analysis	69
4.1	The Design Problem as a Complex Network	70
4.1.1	Ranking Subsystems	72
4.1.2	Ranking Missions	73
4.2	Case Study: AUVs for the Offshore Industry	74
4.2.1	Overview of Deep-Sea Technology	75
4.2.2	AUV Subsystems	77
4.2.3	Subsea Missions in the Offshore Industry	87
4.2.4	Correlation between AUV Subsystems and Offshore Missions .	91
4.2.5	Significance of Subsystems in AUV Design	95
4.2.6	Significance of Missions in AUV Design	96
4.3	Advantage of Flux Metric in Design	97
4.3.1	Comparison to House of Quality	97
4.3.2	Comparison to Other Network Metrics	99
5	Conclusion	101
5.1	Summary	101
5.2	Future Work	102
A	AUV Design Survey	105

List of Figures

1-1	A typical weighted, directed graph	16
1-2	The decomposition of a bipartite graph to a unipartite graph. The nodes α , β , and γ and bipartite edges A, B, C, and D are labeled for clarity.	18
1-3	The communication structure of a hypothetical corporation.	20
2-1	The structure of a corporation where edge weight is assigned according to the authority inherent in a given communication link.	28
2-2	The trigonal grid on which the experiments are conducted. There are four node types arranged so that each node and its three neighbors are distinct.	30
2-3	The stabilization of a conserved flow network following a perturbation. The nodes all have the same weighted indegree and eigenvector centrality score, but node type A has a different flux value than the other nodes.	32
2-4	The stabilization of a Conserved Flow network in which all the nodes have the same weighted outdegree. Node type A has a different flux value than the other nodes.	34
2-5	The stabilization of a network that is composed of four node types with the same weighted indegree but different flux values.	36
2-6	A linear grid composed of three node types.	38

2-7	The stabilization of a linear network after a perturbation at each node type. The nodes all have the same weighted indegree but one node type has a different flux value than the other two types.	39
2-8	A rectangular grid composed of a pattern of five node types.	40
2-9	The stabilization of a rectangular network following a perturbation. Node type A has a different flux value than the other nodes.	41
2-10	The stabilization of a trigonal network after a perturbation. Node types B, C, and D have the same flux value but different flux neighborhoods.	42
2-11	The flux neighborhoods of four nodes. The nodes neighboring along each link are shown in the dashed circles for clarity.	43
3-1	The grid is composed of identical tiles, consisting of a node and its inbound edges, oriented so as to produce varying flux scores for each node type.	49
3-2	Consensus formation in a Synchronization model after a perturbation.	51
3-3	The average state value of Synchronizing nodes following a perturbation at one of four locations.	52
3-4	The size of the infectious population following a spontaneous outbreak at four node types in the Infection model.	58
3-5	The effect of the flux value of the initial site of an epidemic as a function of the infection and removal rates.	59
3-6	The stabilization of a Resource Sharing network after one node receives a perturbation of 250 units.	63
3-7	The maximum decrease in σ achieved in the Resource Sharing model following a perturbation at a node with a flux value of 30% versus a node with a flux value of -50%.	65
3-8	The time at which the ratio of γ to σ_A is the greatest.	66

3-9	The overall improvement in network-wide standard deviation following a a perturbation at a node with a flux value of 30% versus a perturbation at a node with a flux value of -50%.	67
4-1	A bipartite graph modeling the relationships between the subsystems A , B , and C and objectives α and β of a design.	72
4-2	Comparison of power systems [29]	80
4-3	<i>Magnapinna</i> squid photographed by a Shell ROV.	90

List of Tables

- 2.1 Network structures for the Conserved Flow model. 26
- 2.2 An adjacency matrix for a network in which all nodes have the same weighted indegree and eigenvector centrality but one node has a different flux score. 29
- 2.3 An adjacency matrix in which all nodes share the same weighted out-degree but one node has a different flux score. 33
- 2.4 An adjacency matrix for a network in which each node type has a distinct flux value. 36
- 2.5 An adjacency matrix for a linear network in which one node type has a different flux value than the other nodes. 38
- 2.6 An adjacency matrix for a rectangular grid. 40
- 2.7 An adjacency matrix for a network with three nodes that have the same flux value but varying flux neighborhoods. 42

- 3.1 Four network systems from engineering and the social sciences. 46
- 3.2 An adjacency matrix for a network in which the edges represent the fractional distribution of some behavior. 49
- 3.3 A perturbation at a node with a greater flux value leads to a greater steady-state value. 53
- 3.4 An adjacency matrix representing the fractional distribution of a conserved good. 62
- 3.5 The Shared Resource accumulates at nodes with the least flux values. 64

- 4.1 Organizations represented in the AUV design survey. 91

4.2	The correlation between AUV subsystems and offshore oil missions. .	93
4.3	Flux ranking of AUV subsystems.	95
4.4	Flux ranking of offshore missions.	96
4.5	House of Quality ranking of AUV subsystems.	98

Chapter 1

Introduction

The behavior of a complex network arises from the intricate relationships that couple together seeming discrete components, obscuring the relative influence of each node in the larger system. For example, a shock propagating through a network may cause richly complex effects that belie their comparatively simple inception. The consequences of such perturbations vary greatly depending on their point of origin within the network. The influence of a node can therefore be defined in terms of the effects of a shock at that site. The quest to characterize network locations according to these phenomena spans the disciplines of engineering, the physical sciences, mathematics, and the social sciences. This type of network analysis can be applied to challenges as diverse as routing electricity, halting the spread of disease, and designing a robot to explore the ocean. The next sections establish the mathematical structure of the problem, define terminology, and review relevant prior results.

1.1 Motivation: Dynamic Phenomena on Networks

The focus of this project is the influence of individual nodes in a network. This investigation includes ranking the influence of nodes and exploring what significance the location of a perturbation has to the ensuing propagation of changes through the system. The effect of a local shift on the immediate neighborhood may be readily apparent, but the behavior of the system grows increasingly opaque as chain reac-

tions propagate and ripple outward from the initial disturbance. These cascades can also produce feedback effects, which further complicate the system’s behavior. Furthermore, thresholds may exist such that a slight perturbation in one area causes phase shifts across other regions of the network. These nonlinear behaviors together produce the richly complex behavior of the system.

1.1.1 Mathematical Graphs

The qualitative concept of a network is modeled mathematically as a graph. A graph is composed of nodes connected by edges, as shown in Figure 1-1. A graph G is a pair of sets. The edges, E , represent the relationship between any two nodes, which belong to V . Formally, $G = (V, E)$ where $E \subseteq [V]^2$ [20]. If the connected nodes are not distinct, the edge forms a loop. A directed edge is one in which the beginning and ending nodes have been explicitly defined. Two or more edges that share the same beginning and ending nodes are considered to be parallel. An undirected edge is thus equivalent to a set of two antiparallel directed edges.

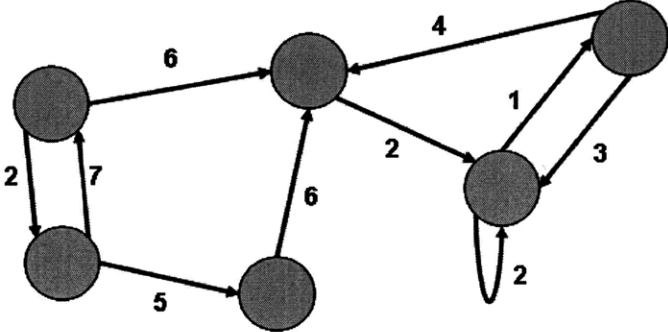


Figure 1-1: A typical weighted, directed graph

Edges may have a weight as well as a direction. The weight of an edge is the strength with which the two nodes are joined when traveling in the direction of the edge. The physical interpretation of an edge’s weight varies among particular applications. For the purposes of this project, it is assumed that the weight of an edge is limited to the positive real numbers. A weight of zero indicates that an edge

does not exist, and a negatively weighted edge would infer the existence of a positive antiparallel edge. Edges with complex weights are outside the scope of this project.

A graph in which an edge may exist between any two nodes is called a unipartite graph. A second important type of graph is a bipartite graph. The nodes of these graphs are divided into two separate and distinct sets. Edges only join nodes of one set to the other, rather than linking nodes of the same set. However, a bipartite graph can always be decomposed to form a unipartite graph composed only of nodes of either set. This transformation is shown in Figure 1-2. Since information is lost during the decomposition, unipartite graphs cannot be converted to bipartite graphs with the same one-to-one correlation.

The applicability of the mathematical construct of a graph to the networks ubiquitous in science and society is readily apparent. Nodes represent self-contained units of the larger system, and edges define how the nodes interact. The scope of this project is limited to simple networks in which the state of a node at any point in time can be described by a single variable. This variable is periodically updated according to the values of any adjacent nodes and the weighting of the edges linking the node to those neighbors. The nodes thus act as autonomous agents with a single state variable.

1.2 Prior Relevant Work in Network Analysis

Common experience indicates that all nodes are not equally influential in a network; every society has dynamic leaders and inconsequential followers. A proper analysis technique should capture the influence conferred upon a node by its location within the network architecture. One measure of the relative significance of a node is its centrality score. The difficulty of formulating a universal definition of “influence” impedes the creation of a network metric to analyze the importance of a node in all possible scenarios. Therefore, several centrality metrics are used concurrently. Each has its own instinctual appeal and appropriately models particular phenomena. In general, many fewer studies have focused on weighted, directed networks than simpler

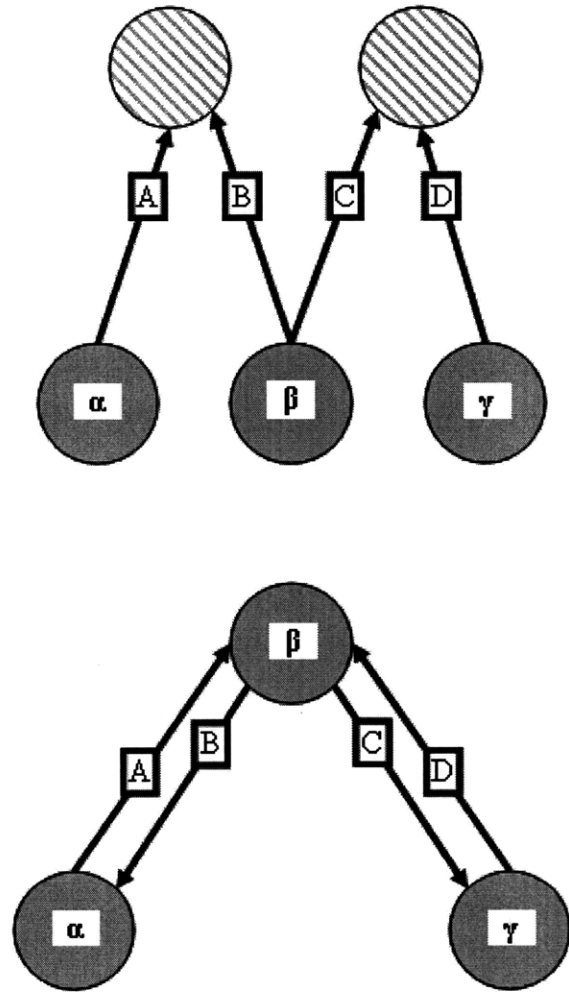


Figure 1-2: The decomposition of a bipartite graph to a unipartite graph. The nodes α , β , and γ and bipartite edges A, B, C, and D are labeled for clarity.

alternatives [51]. This project demonstrates a centrality metric developed for such complex networks by extending the methods previously developed for more simplified, contrived networks.

1.2.1 Degree

The simplest metric used to measure the importance of a node is its degree. In a non-directed network, the degree of a node is simply the number of edges that have that node as either of their terminal points. While rudimentary, degree has been the foundation for some of the most profound discoveries in network science. For example, Watts and Strogatz's description of small-world networks depends on the probabilistic distribution of degree scores among the nodes in the network [75].

Advanced Use of Degree

If the network is weighted, the degree is the sum of the weight of each of those links. The strict definition of degree is inapplicable to a directed network, but closely related metrics have been devised as substitutes [9]. The indegree of a node is the number of edges that end at that node, and the outdegree of the node is the number of edges that emanate from it. Previous research has focused strictly either on indegree or outdegree, failing to unify the two metrics in one form. Therefore, scientists regard the concept of degree as being inapplicable to weighted, directed networks [42].

Secretary Paradox

The simplicity of the degree metric and its intuitive sensibility makes it appealing to many audiences. However, its application can lead to very strange results. For example, consider the corporation represented by the network in 1-3. The nodes of the graph represent various employees, and the edges indicate direct communication among them. The reclusive CEO has no direct contact with her engineers, and only communicates through her secretary. Consequently, the secretary is the only employee with a degree greater than one. This lopsided situation certainly would grant the

secretary a special role in the functioning of the corporation; he would, for example, be aware of any messages passed among the other employees. However, being a conduit for information does not necessarily make the secretary the most influential person at the corporation. While Figure 1-3 depicts an extreme situation, it is typical for network diagrams of more complex and horizontally-integrated corporations to still show the support staff as being much more well-connected than executives. This result is sometimes interpreted as suggesting that secretaries actually are the most powerful individuals in the workplace. However, the apparent paradox could also be used as evidence that a new metric is needed to more accurately measure where the influence lies in a large organization.

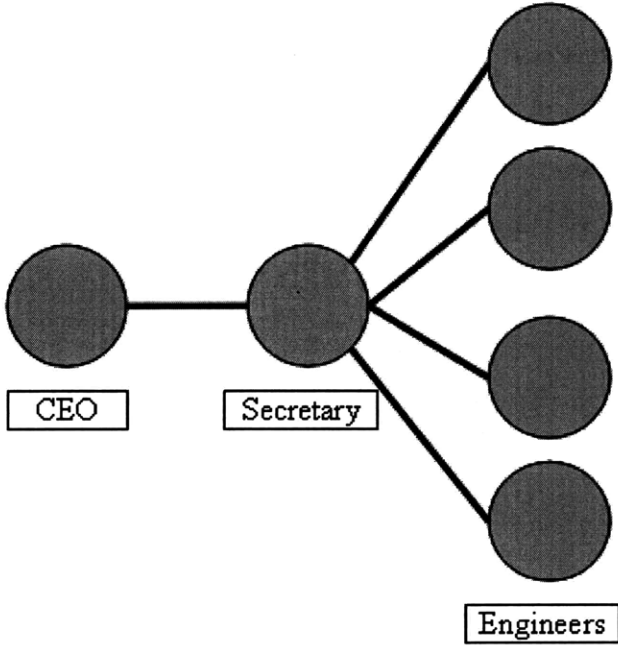


Figure 1-3: The communication structure of a hypothetical corporation.

1.2.2 Betweenness

Freeman [25] refined the betweenness centrality measure, a more sophisticated notion than degree. Betweenness considers the structure of the entire network rather than only a node's immediate neighborhood, the principle limitation of the degree metric.

The betweenness measure of a node is a reflection of how many paths between two other nodes pass through that node.

A node is connected to all other nodes that share its edges and to all nodes that are connected to these immediate neighbors. This recursive definition essentially states that two nodes i and j are connected if and only if a path exists that connects them, where a path is defined as a series of intermediate nodes and edges beginning and terminating at the two nodes of interest. Clearly, many such paths may connect the nodes. Suppose that one such path will be selected at random. There is a finite probability that a particular third node, x , will lay along this path based on the number of applicable paths that happen to pass through x . The betweenness measure of x is the sum of these probabilities for all possible combinations of i and j such that i, j , and x belong to the network and $i \neq j \neq x$.

The betweenness metric was conceived with the goal of gauging the relative significance of nodes in a communication network. A node with a high betweenness score is likely to be used as an intermediary when two other nodes exchange messages. This central node therefore has the ability to regulate the flow of information through the network [25]. However, Borgatti [7] has emphasized that the Freeman's betweenness measure assumes that the packet being passed from node to node is indivisible and does not pass through the same node twice. However, these assumptions are not valid descriptions of many real-world phenomena. For example, it is very likely that a person would tell more than one of their acquaintances a new piece of gossip. The message is therefore not indivisible. Borgatti and his colleague, White, were able to resolve another original shortcoming of the betweenness tool, its inapplicability to directed networks [77]. However, as Thomas [72] has illustrated, Borgatti has not developed a sufficiently universal approach to account for weighted links between nodes, variable behavior of nodes, and other complicating factors. Furthermore, it should be noted that despite its strengths betweenness does not resolve the secretary paradox illustrated in 1-3. The secretary has the only non-zero betweenness score, again incorrectly indicating that he is the most powerful member of the corporation.

1.2.3 Eigenvector Centrality

The most sophisticated measures of centrality are a collection of very similar techniques that depend on eigenvector calculations. Just as the betweenness metric reflects the presumption that a message is being relayed between intermediaries, eigenvector methods are derived from the qualitative notion of popularity in a social network. This technique has been introduced by Phillip Bonacich [4], and is frequently referred to by his name.

The premise underlying the Bonacich metric is that when calculating power in a social network not all relationships are equal. It is more valuable to be associated with individuals who are popular than it is to be linked to individuals who are otherwise isolated. This concept is intuitive; in any society an individual who is popular with popular people has disproportionate influence. However, this definition of centrality depends on circular reasoning: a person's popularity is judged by the popularity of his friends, but their popularity depends on the popularity of their own associates, including him. If the popularity of each of n nodes in a network is a value in the vector \mathbf{P} of length n , then the influence structure can be stated mathematically as

$$c\mathbf{P} = \mathbf{A}\mathbf{P} \quad (1.1)$$

The $n \times n$ matrix \mathbf{A} in Equation 1.1 is an adjacency matrix, a convenient mathematical form for representing networks. Each value (i, j) of the matrix corresponds to the weight of the link between nodes i and j . For a directed network the convention is adopted that the value listed is the strength of the link from node j to node i . When the network is undirected, the adjacency matrix would be symmetrical across the diagonal.

Equation 1.1 is an eigenvector equation, where the constant c is a particular eigenvalue and \mathbf{P} is the corresponding eigenvector. Traditionally the eigenvector corresponding to the largest eigenvalue is considered to be the Bonacich centrality scores of the nodes of the network. For a symmetric adjacency matrix, i.e. an undirected network, the centrality scores are guaranteed to be non-negative.

Bonacich’s elegant solution to the challenge of self-defined popularity via network connections has proven to be immensely popular. For example, it has been used to analyze co-authorship networks and judged superior to other centrality measures [42]. Recent research has also demonstrated that eigenvector centrality measures the criticality of a node in a dynamic network, i.e. which nodes’ removal would be most disruptive to the structure of the network [54]. Even the famous PageRank algorithm that drives the Google search engine is based on the same principle as eigenvector centrality [8].

Eigenvector methods are more versatile than degree and betweenness because Bonacich’s technique is applicable to weighted, directed networks [5]. However, as Bonacich himself has emphasized [6], caution must be exercised when applying eigenvector techniques to these special graphs. For example, a directed relationship extended from an otherwise isolated member of a social network does not give any increase to its recipient’s eigenvector centrality score. However, in many real-world networks members have an intrinsic value independent of their relationship to their peers. To compensate for this specific disparity between reality and the network tool, Bonacich has introduced further refinements to his algorithm [6]. Other researchers have also tweaked Bonacich’s approach to broaden its applicability. For example, in order to compare eigenvector centrality scores across different graphs, it has been demonstrated that normalization is necessary to compensate for the varying number of nodes in each graph [57]. This constant modification demonstrates that no network metric is universally applicable, and that each quantitative model must be coupled to a qualitative understanding of the applicable scenario or else it risks becoming devoid of meaning.

1.3 Problem Statement: Identifying Influential Nodes and Their Role in Engineering Design

The goal of this project is to develop a novel network analysis metric that successfully identifies influential nodes. This technique is meant to differentiate among nodes that other methods regard as identical but whose behavior is distinct. Special attention is paid to weighted, directed networks due to their applicability to many challenges.

Influence is the ability to exert one's will onto others. Therefore, the state value of a more influential node should have a greater impact on the state of the entire system than the value of a less influential node. This criterion will be the benchmark for a series of experiments to test if the novel metric adequately predicts the influence of a node. Networked systems at a state of equilibrium will be perturbed and the influence of the site of the disturbance will be measured by the ensuing return to equilibrium. The network architecture will remain constant during all of the experiments; the focus of the study is the response of the system to changing state values rather than network links.

The concept of influence within a network will then be expanded to the novel application of engineering design. The use of a network model to represent a design challenge is demonstrated through a case study, the design of autonomous underwater vehicles for the offshore petrochemical industry. This project thus seeks to expand both the capabilities and applications of network analysis measures.

Chapter 2

Introducing the Flux Metric

Introduction

The need exists for a technique to measure the influence of nodes in weighted, directed networks. This new method should be independent of a specific model framework and applicable to a wide variety of scenarios. Furthermore, it should serve as a complement to eigenvector centrality and other existing network metrics by registering nodal characteristics that elude these measures. The solution to this challenge considered in this project is the flux metric, a simple mathematical concept with profound implications.

The flux metric is a novel measure of nodal influence inspired by fluid mechanics. This method offers insight into networks that other centrality measures fail to provide on a wide variety of network architectures. This chapter uses a Conserved Flow model to demonstrate the significance and appropriate interpretation of the flux metric, such as the proper comparison of two nodes' flux scores and the concept of flux neighborhoods, through a series of experiments on abstract networks. Table 2.1 is an overview of the network structures used in this chapter. These experiments provide the basis for the application of the flux metric to a wide variety of real-world scenarios.

Network Geometry	Number of Flux Values	Indegree of Nodes	Flux Neighborhoods of Nodes with Equal Flux
Trigonal	2	Identical	Identical
Trigonal	2	Different	Identical
Trigonal	4	Identical	N/A
Linear	2	Identical	Identical
Rectangular	2	Identical	Identical
Trigonal	2	Identical	Different

Table 2.1: Network structures for the Conserved Flow model.

2.1 Concept of Flux

In a weighted, directed network each node may have edges directed from and toward the node. Suppose that ϕ is a function of the sum of the weights of either the outgoing or incoming edges. The flux of a node is defined as

$$flux = \phi(outdegree) - \phi(indegree) \quad (2.1)$$

In the simplest and most common case considered in this project, ϕ is simply the identity function and the flux is the difference between the weighted outdegree and indegree. Note that in this case the flux values of all the nodes in a network will always sum to zero, because the outdegree of a node is accounted for by the indegrees of its neighbors. If the edges represent the rate of flow through network conduits, the flux is a measure of the rate at which a material accumulates or dissipates at each node. The flux metric is thus consistent with the property of the same name from the physical sciences, i.e. the net flow through a control surface. Surprisingly, the transfer of this simple and ubiquitous concept from fluid mechanics, physics, and other scientific disciplines to network analysis has almost never been considered previously. Taylor has incorporated the net difference between incoming and outgoing edges in unweighted, directed networks as part of a much more complicated algorithm designed to identify influence clusters in networks [71]. However, this technique has not been explored further by other researchers. It is possible that their lack of interest has been due to the belief that Taylor's method is unnecessary in light of Bonacich's later work.

Alternatively, the overall complexity of Taylor’s algorithm may have obscured the utility of the simple, yet powerful, concept of considering the net difference between the weighted outdegree and indegree of a node.

2.2 Comparison of Flux to Previous Metrics

The advantage of the flux metric is that it shares degree centrality’s independence from a particular paradigm or case study, but is applicable to weighted and directed networks. The flux metric provides insight into the relative influence of nodes that is distinct from the conclusions of the other network metrics. This project focuses on the influence of a node in the context of dynamic shock propagation and static dominance in a social or design network. This phenomenon and the manner in which the flux metric registers the consequences of a perturbation at a node can be demonstrated empirically.

2.2.1 Flux and the Secretary Paradox

The flux metric permits the analysis of more realistic models of complex systems than are possible using other measures of centrality. For example, the flux metric resolves the Secretary Paradox introduced in Section 1.2.1. This apparent paradox arose because an overly simplified yet typical representation of a corporation’s structure used unweighted and undirected edges that made it appear as though the secretary were more influential than the CEO. A more accurate depiction of the corporation can be made using weighted and directed links as shown in Figure 2-1.

The strength of the edges represents the level of authority associated with a command issued by each actor in that particular relationship. The flux of the CEO is 4, the flux of each engineer is 1, and the flux of the secretary is -8. The ranking of the nodes according to their flux is a much more reasonable result than the conclusion reached by simply examining the degree of each node. That approach yields that the CEO and her engineers are peers and subordinate to their secretary. The flux metric also provides a more accurate ranking than eigenvector centrality. While

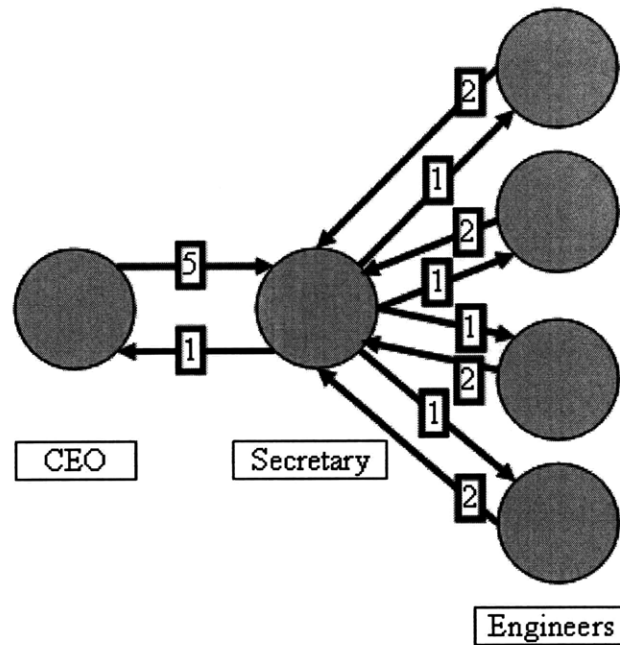


Figure 2-1: The structure of a corporation where edge weight is assigned according to the authority inherent in a given communication link.

Bonacich's method recognizes that the secretary is the least important member of the corporation, it ranks the CEO and engineers as peers because they all have the same weighted indegree. Thus, the Secretary Paradox demonstrates both the great value inherent in a network analysis tool that acknowledges the variability among the relationships comprising the same complex system and the ability of the flux metric to provide a more refined measure of the influence of a node than existing centrality measures.

The application of the flux metric to the Secretary Paradox also illustrates a unique characteristic of this technique. It is possible for a node to have a negative flux value, something that none of the traditional centrality measures permit. This occurrence could be interpreted as stating that the node is a net follower rather than a net leader in its relationships. The node is therefore more likely to be influenced by its immediate neighbors than it is to affect their state.

	A	B	C	D	Flux
A	0	3	3	3	-6
B	1	0	4	4	2
C	1	4	0	4	2
D	1	4	4	0	2

Table 2.2: An adjacency matrix for a network in which all nodes have the same weighted indegree and eigenvector centrality but one node has a different flux score.

2.2.2 Network Models

The Secretary Paradox concisely demonstrates the ability of the flux metric to appropriately rank social influence in an organization. However, before flux can be applied to engineering networks it is necessary to model and quantify the significance of a node’s flux value in dynamic processes. A special graph is used as the framework for this process. The network is composed of 160 nodes arranged in a tiled trigonal pattern, an example of which is shown in Figure 2-2. Note that each link shown between the nodes in the figure is actually a set of two oppositely directed weighted edges. The nodes are categorized into one of four types, and the nodes of each types behave identically. The experiments are facilitated by the fact that each node and its three neighbors are of separate types. Note that the nodes along the “top” and “bottom” and along the “right” and “left” of the network are joined, giving the network a toroidal shape. This structure eliminates the possibility of boundary effects that could complicate the output of the experiments and confuse the interpretation of the results.

2.2.3 Unique Information Revealed by the Flux Metric

The 160-node trigonal network is the framework for a Conserved Flow model. The weightings of the edges between the four node types are assigned as shown in the adjacency matrix in Table 2.2.

The table shows that each node has the same weighted indegree of 9. After forming an adjacency matrix to represent the entire 160-node network and applying Equation 1.1, it can be shown that every node shares the same eigenvector centrality score of

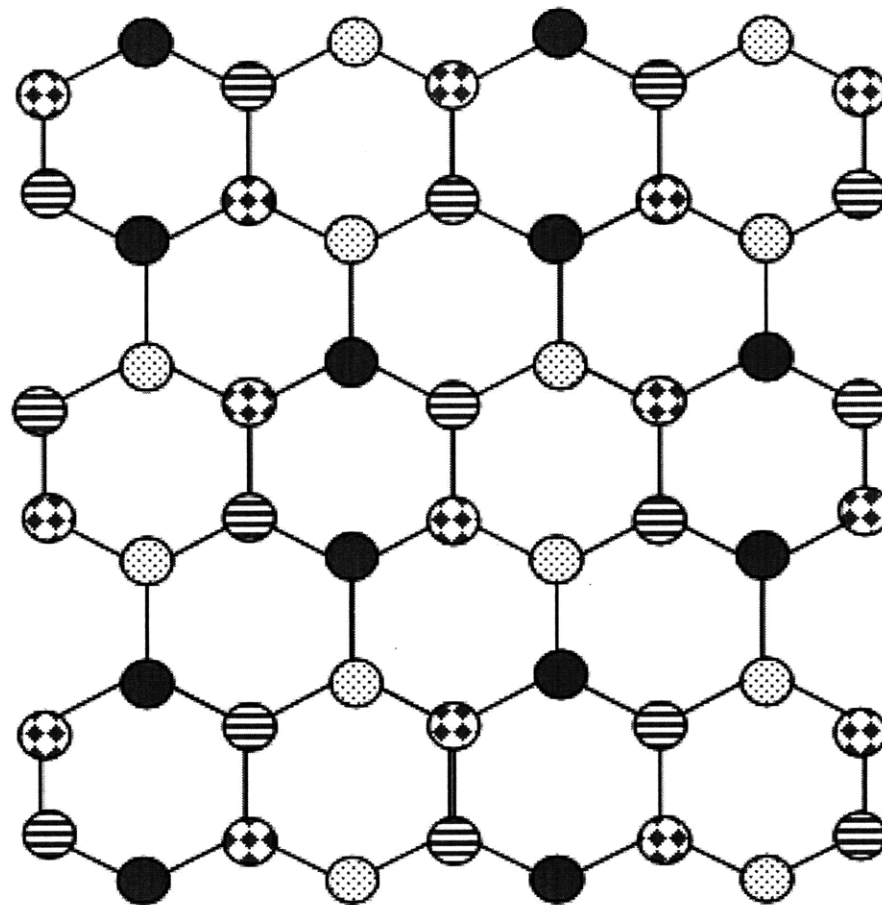


Figure 2-2: The trigonal grid on which the experiments are conducted. There are four node types arranged so that each node and its three neighbors are distinct.

0.0791. However, node type A has a flux value of -6 while the other nodes have a flux value of 2.

The weighting of each directed edge represents the maximum flow per unit time through that link. The state value of each node represents the amount of material present at that location at a given time. This model is well-suited to modeling the flow of a conserved entity such as electricity or water. To simulate the effect of a perturbation on the system while it is at equilibrium, the state values of all but one node initially are set to zero. A shock is then induced by setting the initial value of the perturbation site to 250.

The Conserved Flow model simulates simultaneous exchanges between the nodes by updating the state value of each node once per time step. If the state value of a node is greater than the amount of material that can exit the node in one time step, i.e. the node's weighted outdegree, then the maximum amount permissible passes out of the node and through each conduit to the node's neighbors. If the node's current state value is less than this maximum flow capacity, then the entire amount present is transferred to its neighbors proportionately to the weighting of each outgoing edge.

The flow between nodes permits the network to reach a new equilibrium after the initial perturbation. However, the path to this steady state depends on the type of node at which the perturbation begins. In the Conserved Flow model a shock at a node with a greater flux value is expected to lead to a state of equilibrium more quickly because the node has greater influence over the system than other nodes do. The level of consensus present in the system at any time is quantified by the instantaneous standard deviation of the state values of all the nodes in the network. This quantity at time t following a perturbation at a type A node is designated $\sigma_A(t)$. The level of consensus at any point in time is inversely related to the value of σ . Note that as time approaches infinity the nodes do not reach uniform nor constant values due to the structure of the graph. Therefore, a universal consensus at which σ would equal zero cannot be achieved. However, σ does reach a steady state. Figure 2-3 shows the value of σ for 250 time steps following a perturbation at one of the four node types. The significant features of Figure 2-3 are the relative positions of

the σ plots rather than the nuances of those functions. The abrupt shifts, dips, and other nonlinear details are evidence of the complex behavior that can arise in network functions but the overall influence of a node is reflected by the the value of σ at a given point in time following a perturbation at that site.

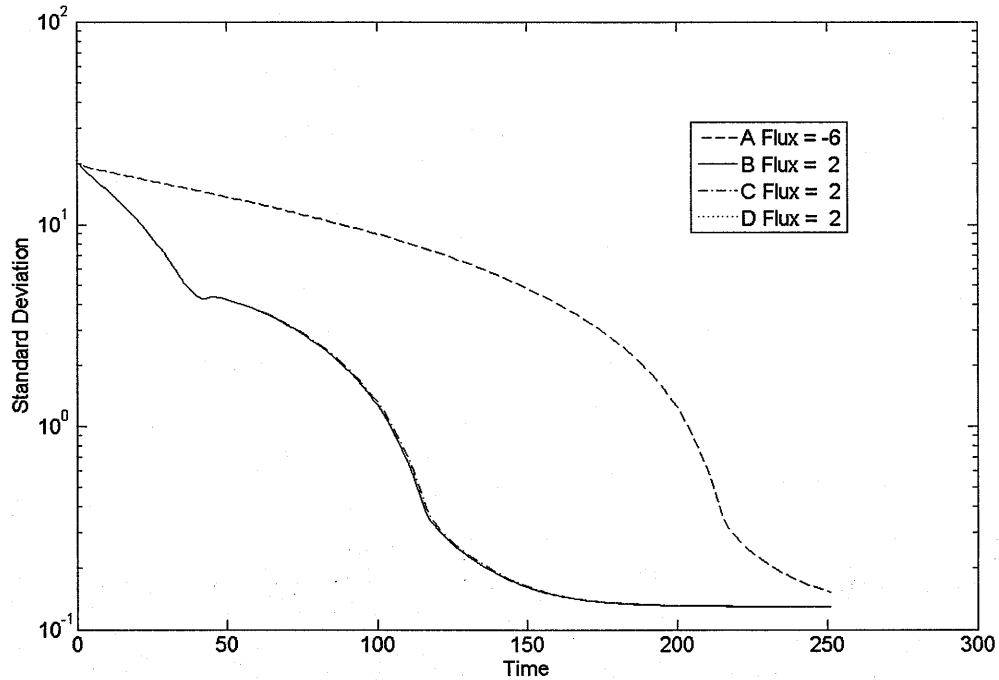


Figure 2-3: The stabilization of a conserved flow network following a perturbation. The nodes all have the same weighted indegree and eigenvector centrality score, but node type A has a different flux value than the other nodes.

The key result depicted in Figure 2-3 is that after a perturbation at one of the lower flux nodes the level of consensus in the system, as measured by the inverse of σ , is 10 times less than if the perturbation originates at a node with a greater flux value. Thus nodes that both the eigenvector centrality and weighted indegree metrics regard as identical have dramatically different properties in terms of their influence on consensus formation.

It is clear that a surge would initially dissipate more quickly from a site with a greater net outflow. However, the surprising result is that this initial advantage con-

	A	B	C	D	Flux
A	0	1	1	1	6
B	3	0	4	4	-2
C	3	4	0	4	-2
D	3	4	4	0	-2

Table 2.3: An adjacency matrix in which all nodes share the same weighted outdegree but one node has a different flux score.

tinues for nearly the entire duration of the simulation. Due to the regular architecture of the network depicted in Figure 2-2, a perturbation at any location immediately will encounter all of the other node types after exiting the site of origin. Thus, there is no obvious reason why a perturbation would dissipate more quickly across the entire network, rather than just the immediate site of the perturbation, after beginning at a high-flux node.

Furthermore, consensus forms more quickly when the initial disturbance to equilibrium occurs at a node with a lower flux value than at a node with a higher flux value even if all the nodes have the same weighted outdegree. Such a network can be constructed by reversing the direction of each edge in the adjacency matrix in Table 2.3 to form the adjacency matrix in Table 2.3. Note that it is impossible for nodes to have the same eigenvector centrality unless they also have the same weighted indegree or for two nodes to have the same indegree and outdegree but different flux values.

Each node type in Table 2.3 has the same weighted outdegree, 9, but node type A has a flux value of 6 while every other node type has a flux value of -2. The values of σ over time following a perturbation at each node type are shown in Figure 2-4. It is very clear that at any point in time the value of σ is less if the perturbation originates at a higher-flux node than at a node with a lower flux value.

A very significant result in Figure 2-3 and Figure 2-4 is that the behavior of σ is the same after perturbations at the three node types with identical flux scores. Very slight deviations from this trend are due to rounding effects that occur when the state values of the nodes are very small. The observation that the nodes with identical flux values exhibit the same behavior supports the notion that the flux metric

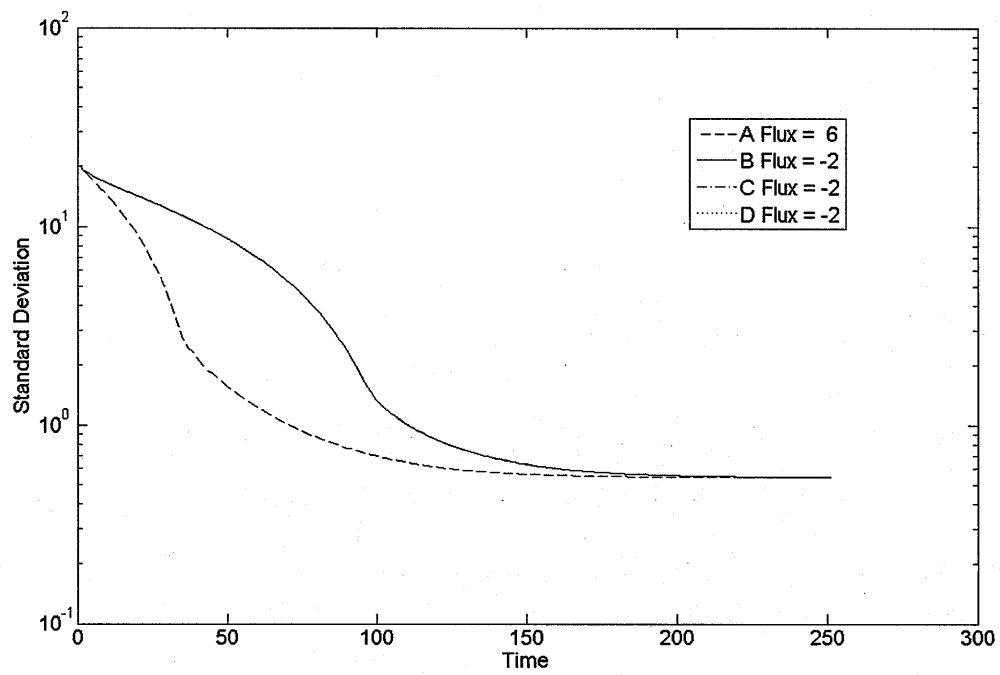


Figure 2-4: The stabilization of a Conserved Flow network in which all the nodes have the same weighted outdegree. Node type A has a different flux value than the other nodes.

can successfully categorize nodes according to their influence on consensus forming. Therefore, the flux metric not only reveals information about the nodes that escapes other centrality scores but also appears to fully capture the nodes' influence in this scenario.

2.3 Exploring the Flux Metric

The ability of the flux metric to indicate patterns of consensus formation following perturbations at different nodes is clear. However, to apply flux to a wide variety of real-world situations it is necessary to understand the full extent of the metric's meaning and significance. These properties can be demonstrated by variants of the Conserved Flow experiment.

2.3.1 Comparing Flux Values

The original Conserved Flow experiment demonstrates that at any point in time σ will be less following a perturbation at a node with a greater flux value than at another node. However, further experiments are required to better understand the relationship between this gap the relative difference in the flux values of the nodes. For example, simulations not shown here confirm that in a Conserved Flow scenario multiplying the weighting of each edge by a constant has no effect on the behavior of σ . However, non-proportionate changes in the flux value of each node have more interesting effects.

The adjacency matrix in Table 2.4 is deliberately arranged so that each node has a different flux value. Note that each node in the trigonal network still has the same eigenvector centrality score and a weighted indegree of 12. This structure enables an analysis of the behavior of σ as a function of the flux value. The result of a Conserved Flow experiment in this scenario is shown in Figure 2-5.

It is clear that the rate of consensus formation shown in Figure 2-5 is not proportional to the flux value of the node at which the perturbation began. For example, a perturbation at node type C, which has a flux value of 4, causes consensus formation

	A	B	C	D	Flux
A	0	4	4	4	-8
B	2	0	3	7	8
C	1	6	0	5	-4
D	1	10	1	0	4

Table 2.4: An adjacency matrix for a network in which each node type has a distinct flux value.

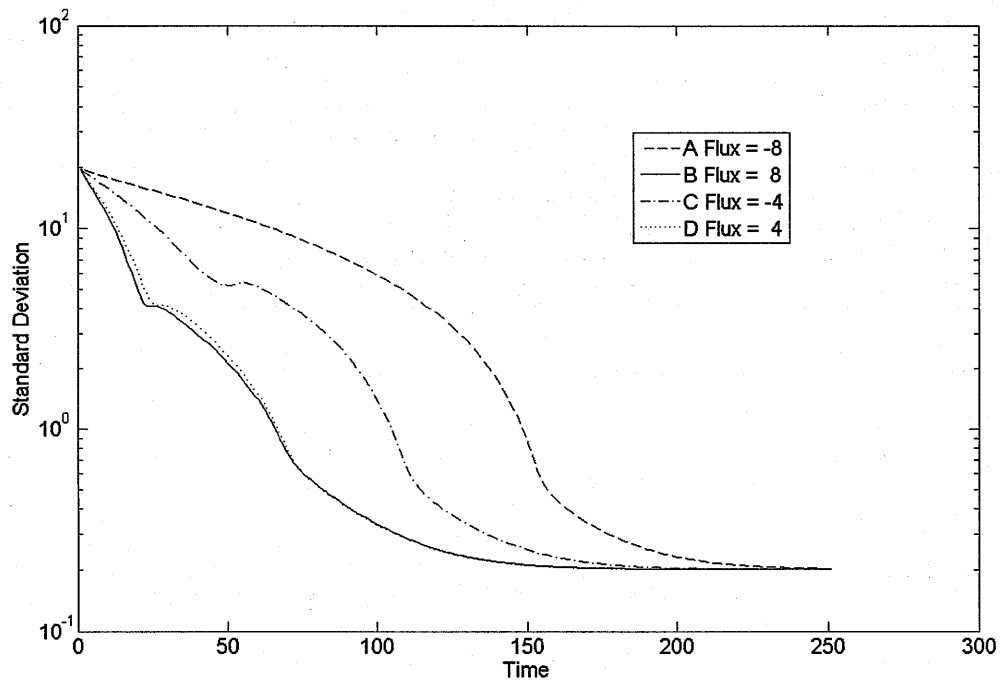


Figure 2-5: The stabilization of a network that is composed of four node types with the same weighted indegree but different flux values.

approximately intermediate to the behavior following a perturbation at the nodes with flux scores of -8 and 8. However, the flux values of the nodes still predict the ranking of the rate at which consensus is achieved following a perturbation at that node type. Thus, flux is a strong indicator of relative influence in a network by means of consensus forming but two flux scores can only be compared usefully in terms of having a greater, lesser, or equal value rather than their actual arithmetic difference. Correlating σ to a particular flux value beyond this simple relationship is extremely difficult because the flux value of every other node in the network, their arrangement, and the specific nature of the dynamics between them can all potentially change the effect of a perturbation occurring at the node.

2.3.2 Flux in Graphs of Varying Geometry

A major motivation for the development of the flux metric is the need for a measure of nodal influence that is applicable to many different network structures. Therefore, it is necessary to demonstrate that the properties of the metric do not vary when it is applied to different graphs. This premise can be tested by modifying the network on which the Conserved Flow simulation is conducted.

A simple modification to the trigonal grid introduced in Section 2.2.2 is to increase the number of nodes in the graph from 160 to 6,400 while preserving the toroidal structure. This alteration tests if the correlation between σ and the relative flux values of two nodes is a fluke occurrence on a 160-node graph or if it also applies to relative large scale networks. The results of experiments not shown here indicate that the declared properties of the flux metric are indeed independent of the size of the network under investigation.

Flux in Linear Graphs

The flux value of a node also indicates its influence during Conserved Flow on non-trigonal graphs. For example, consider the simple network formed by a linear chain of three different node types. In a fashion similar to the architecture of the trigonal

	A	B	C	Flux
A	0	3	3	-4
B	1	0	5	2
C	1	5	0	2

Table 2.5: An adjacency matrix for a linear network in which one node type has a different flux value than the other nodes.

network, the node types are assigned so that each node and its two neighbors are separate types as shown in Figure 2-6. The ends of the graph link together to form a loop to avoid boundary effects. The weighting of each directed edge is shown in Table 2.5.



Figure 2-6: A linear grid composed of three node types.

The nodes all have the same weighted indegree and eigenvector centrality but node type A has a flux value of -4 while node types B and C have flux values of 2. A Conserved Flow experiment beginning with a perturbation of 100 is executed on this 300-node network, in which each node initially has an equilibrium value of zero. The behavior of σ for 750 iterations following a perturbation at each node type is shown in Figure 2-7.

Figure 2-7 shows that σ is smaller following a perturbation at a node with a lesser flux value than after a change at a node with a greater flux value. Furthermore, perturbations at nodes with the same flux values result in identical displays of consensus formation. Therefore, the flux metric is fully applicable to simple linear graphs.

Flux in Rectangular Graphs

The applicability of the flux metric to the analysis of more complex graphs than the trigonal grid can be demonstrated using a rectangular graph. This graph is composed

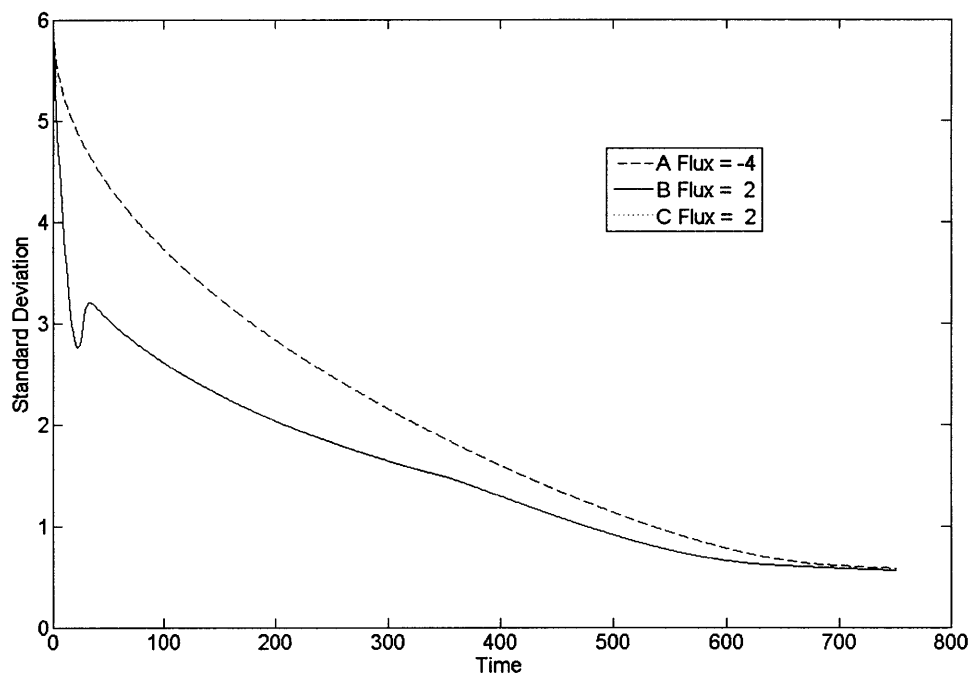


Figure 2-7: The stabilization of a linear network after a perturbation at each node type. The nodes all have the same weighted indegree but one node type has a different flux value than the other two types.

	A	B	C	D	E	Flux
A	0	4	4	4	4	-12
B	1	0	5	5	5	3
C	1	5	0	5	5	3
D	1	5	5	0	5	3
E	1	5	5	5	0	3

Table 2.6: An adjacency matrix for a rectangular grid.

of a repeating pattern of five different types of nodes, the structure of which is shown in Figure 2-8. The adjacency matrix for the five node types is shown in Table 2.6.

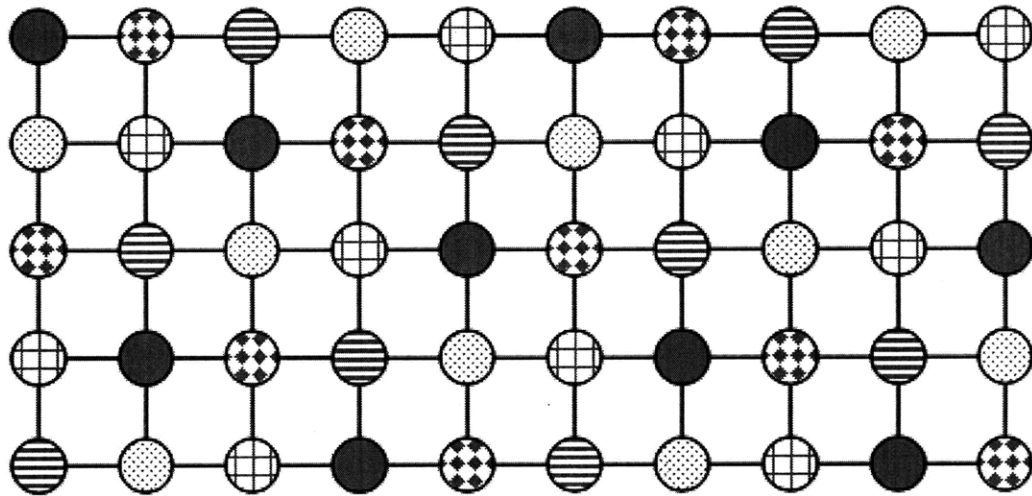


Figure 2-8: A rectangular grid composed of a pattern of five node types.

In this scenario nodes of type A have flux values of -12 while all other nodes have flux values of 3. A Conserved Flow begins with a perturbation of 500 on the 200-node network, the members of which have an initial equilibrium state value of zero. The ensuing consensus formation is shown in Figure 2-9.

The perturbation dissipates throughout the network more quickly if it begins at a node with a lesser flux value than a node with a greater flux value. In addition, σ is identical following a disturbance at any of the equally-valued nodes. Therefore, the principles of flux analysis hold true on more complex graphs than the linear or trigonal networks.

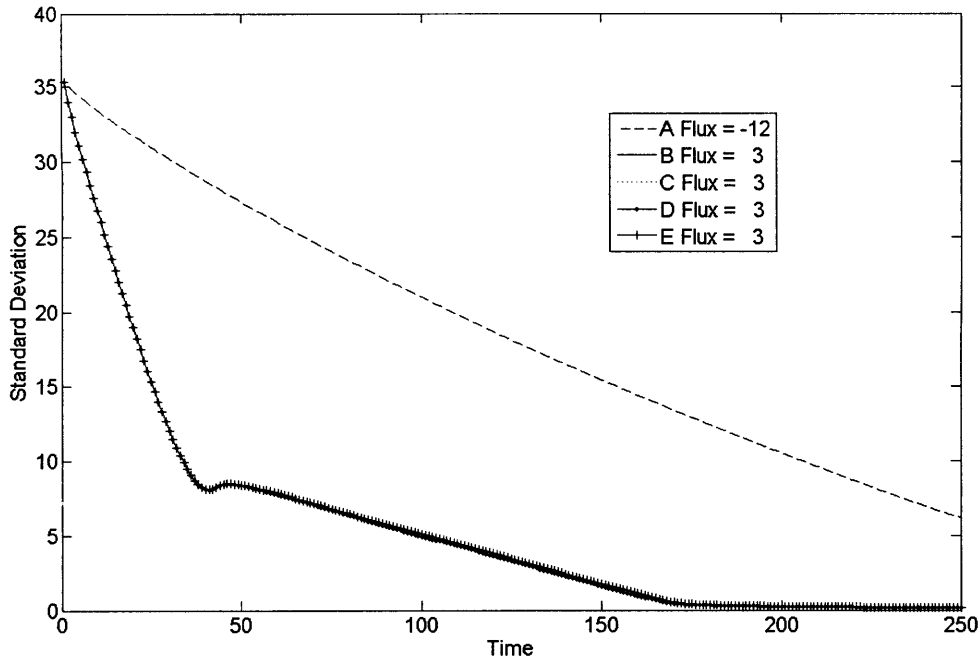


Figure 2-9: The stabilization of a rectangular network following a perturbation. Node type A has a different flux value than the other nodes.

2.3.3 Flux Neighborhoods

Flux, as defined thus far, is actually not the sole determinant of consensus-forming behavior following a perturbation. It is possible for two nodes in a network to have the same flux value yet for the change in network-wide standard deviation, σ , to vary following a perturbation at each node. However, flux remains the strongest metric for ranking the overall influence of the nodes.

A variation of the Conserved Flow experiment on the 160-node network demonstrates this phenomenon and justifies the claim that flux remains significant despite this limitation. The adjacency matrix describing the relationships between the four node types is shown in Table 2.7. Note that node type A has a flux value of -3 while the other node types have flux values of 1. The result of perturbations occurring at each node type are shown in Figure 2-10.

Figure 2-10 clearly demonstrates that perturbations at nodes with the same flux values do not necessarily result in the same consensus-forming behavior, as has been

	A	B	C	D	Flux
A	0	1	2	3	-3
B	1	0	3	2	1
C	1	3	0	2	1
D	1	3	2	0	1

Table 2.7: An adjacency matrix for a network with three nodes that have the same flux value but varying flux neighborhoods.

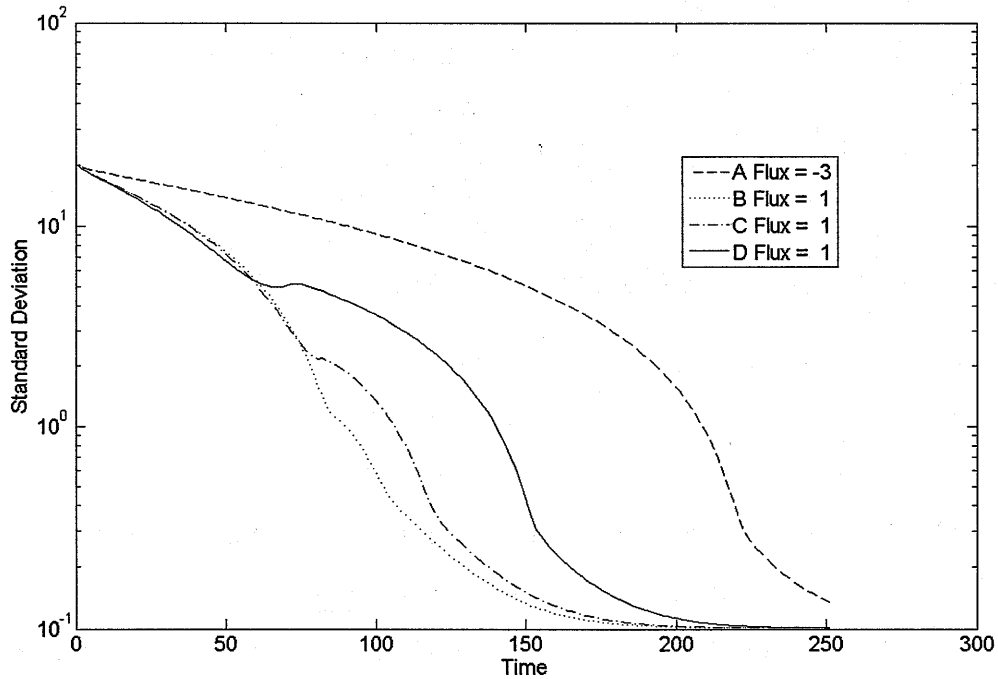


Figure 2-10: The stabilization of a trigonal network after a perturbation. Node types B, C, and D have the same flux value but different flux neighborhoods.

true in all previous simulations. The cause of this discrepancy lies in each node's flux neighborhood, the net flow between the node and its individual neighbors. The flux neighborhood is the arrangement and value of $\phi(\text{outdegree})$ minus $\phi(\text{indegree})$ along each collection of edges between a node and a particular neighbor. For example, the flux neighborhoods for the scenario in Table 2.7 are illustrated in Figure 2-11.

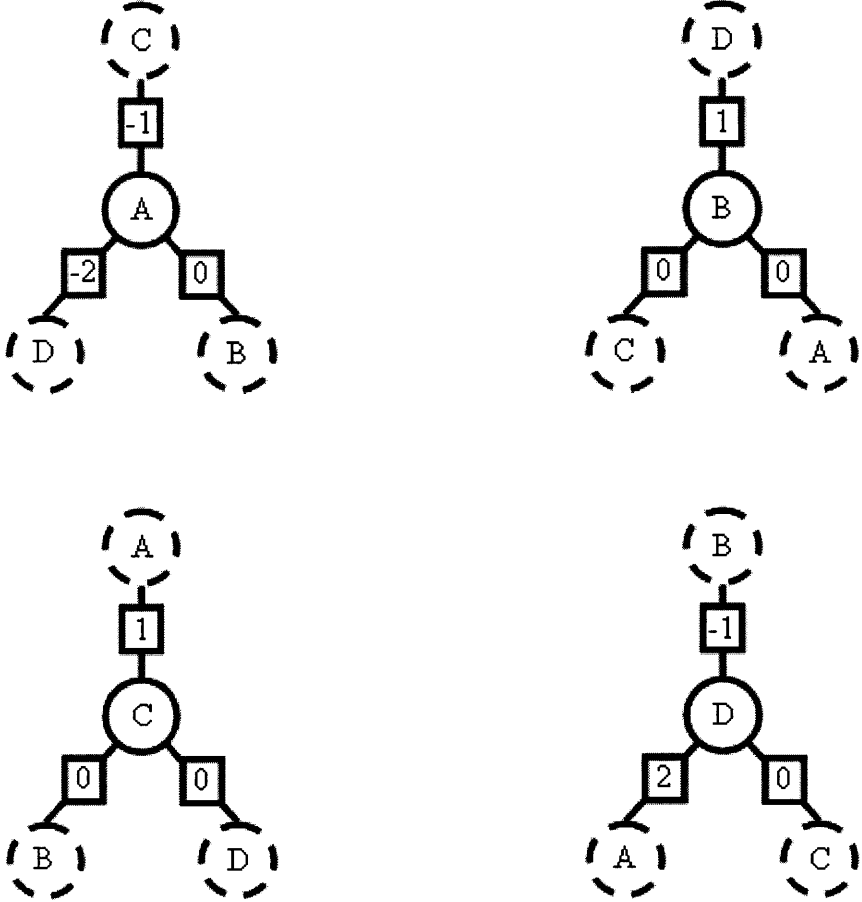


Figure 2-11: The flux neighborhoods of four nodes. The nodes neighboring along each link are shown in the dashed circles for clarity.

Flux neighborhoods are defined independently of orientation, i.e. two nodes have the same flux neighborhoods if one node's neighborhood values are an axial rotation of the other node's neighborhood values. Node types B and C in Figure 2-11 thus have the same flux neighborhood. Node type D has a separate flux neighborhood even though this node has the same flux value as node types B and C.

Experiments not shown here indicate that unless all nodes in a network that have

the same flux value also share the same flux neighborhood that a perturbation at any of these nodes will not produce the same σ behavior. Among nodes with the same flux score, σ may rise or fall as compared to the results of another node with the same flux value. However, the network-wide standard deviation following a perturbation at any node of a particular flux value is still bounded within the standard deviation produced by a perturbation at nodes of greater or lesser flux scores. If two or more nodes with the same flux value do have the same flux neighborhood, the difference in σ following a perturbation at each of these nodes is smaller than the difference between either of their σ values and that of a node with the same flux value but a different flux neighborhood. These phenomena are all present in Figure 2-10 and are also observed following similar scenarios using other network structures.

It is not clear how to meaningfully characterize flux neighborhoods but the flux score is still a relevant metric despite this lack of knowledge. The influence of the nodes can consistently and reliably be ranked according to the flux values of the nodes. The only uncertainty is how to rank the outfall of perturbations at two nodes with the same flux values. Furthermore, analysis based on weighted indegree or eigenvector centrality would have concluded that all of the node types listed in Table 2-10 were identical, so the flux metric still yields greater insight than any major centrality score despite its limitations.

Conclusion

The flux metric, which is defined according to Equation 2.1, accurately indicates the relative influence of nodes in the Conserved Flow network. Influence in this scenario is defined as the ability for a perturbation of material to be dispersed through the network and thereby establish a new steady state. The concept of flux neighborhoods adds further detail to the interpretation of flux values. Flux is able to differentiate among nodes that other network metrics regard as identical. These conclusions can be demonstrated on trigonal, linear, and rectangular networks of various sizes.

Chapter 3

Applications of the Flux Metric

Introduction

The interpretation of the flux metric as a measure of nodal influence has been explored and verified on regular, tessellated network structures. These models may appear overly abstract for the ensuing results to be readily applicable to real-world challenges. However, three significant complex systems are described in detail in this chapter that may be represented as dynamic agents operating on such networks.

The Conserved Flow model presented in Chapter 2 is directly applicable to the passage of electricity through a power grid, water in a plumbing system, or supplies through distribution centers. The key characteristics of these systems are the conserved, material goods that they transport and the limited capacity of the conduits involved. In this chapter three more models of network dynamics on the trigonal network structure are discussed. The synchronization of data or opinions, the spread of disease, and the allocation of resources in a community are all forms of network stabilization following a perturbation. Each model presented here simulates one or more such pertinent systems from engineering and the social sciences, and previous research manifests the importance of these challenges to their respective disciplines. The models are summarized in Table 3.1. Special attention is paid to those networks on which nodes have uniform weighted indegree values, and hence equal eigenvector centrality scores, because these cases represent the greatest opportunities for the flux

Model	Conserved Material	State Variable	State Value Update Method
Conserved Flow	Yes	Continuous	Passed through limited-capacity channels to neighbor
Synchronization	No	Continuous	Weighted average of neighbors' values
Infection	No	Discrete	Weighted probability of contraction from neighbors
Resource Sharing	Yes	Continuous	Proportionately distributed to neighbors

Table 3.1: Four network systems from engineering and the social sciences.

metric to make a new contribution to network studies.

3.1 Synchronization

A distributed system may reach a shared opinion without the benefit of a central authority. Consensus algorithms that govern the local behavior of agents to achieve this global goal have been the focus of much previous work. DeGroot, for example, has developed a model to represent a committee debate on a topic represented by a continuous variable [18].

The correlation between the structure of a network and consensus-forming behavior has also been studied [21]. For example, the rate at which values converge may be characterized by the eigenvalues of the network's Laplacian [67]. The eigenvalues of the Laplacian, a representation of a network similar to an adjacency matrix, may also indicate the type of cooperation between robots distributed in a network [23] [65]. Recall from Section 1.2.3 that the first eigenvector of the adjacency matrix contains the Bonacich centrality scores of the nodes. Eigenvector analysis can thus describe the behavior of a network in rich detail. However, few studies have investigated the influence of individual nodes in the network during Synchronization.

A novel idea that one individual introduces to a community may diffuse through the entire network until it has affected the perspective of every member, despite their distant location from the fad's point of origin. A Synchronization model can

illustrate how this process depends on the properties of the node at which it begins. This insight would be relevant to the study of not only social phenomena but also engineered systems such as sensor networks.

Distributed sensor networks take advantage of a trend towards manufacturing sensors that are more prolific but lower in cost and quality [79]. The task of sensing is distributed across many units, and the community consensus value is reported to the user via a designated gateway sensor. The advantage of such an array is its resiliency to failure, particularly in hazardous environments such as the ocean where each individual unit has a low probability of survival. Sparsely distributed sensors may only have sufficient power to broadcast to a small number of neighbors. Information may leave the network only when it is relayed to the gateway node. This type of acoustic underwater network has been proposed by Akyildiz et al. [1]. It is important to consider how an anomaly due to a malfunctioning sensor would affect the final data report. Analogously to the effect of an innovation appearing in the social network, a perturbation caused by a faulty sensor could ripple through the system with a particular speed dependent on its point of origin.

3.1.1 Relevant Previous Work

The core concepts of the Synchronization model are cooperation and the flow of information in a network. This phenomenon is an extraordinarily popular topic in network science and a full review of the relevant previous work is beyond the scope of this project. Communication and opinion formation are considered so important because they lie at the heart of social interactions and the effort to quantify human relationships.

It has been assumed that an individual's social associates affect their opinions, and that stronger relationships correlate with a greater degree of trust and hence mutual influence. Yardi and Bruckman have used the Facebook social networking website to test whether trust, as measured by the exchange of videos and games, actually correlates to the strength of a relationship as indicated by shared friends, groups, and other Facebook parameters [78]. The web of relationships that define most social

systems are not as conveniently defined as Facebook profiles. Therefore, known formal associations must be used as proxies for the more ethereal relationships that define human communities. For example, in one popular case study, nodes represent families in Renaissance-era Florence and edges represent alliances in the form of marriages [74]. Studies of the United States Congress have used bill cosponsorships to map the relationships among lawmakers [24]. Issues of trust and communication are also prominent in studies of overlapping corporate boards of directors and the networks that exist among national governments that govern international relations [41] [44].

After the network structure has been established the ability of a node to influence other members of the community or to establish a shared opinion may not be clear. For example, Moreau recognizes a critical level of communication that permits network-wide coordination [49]. While it is intuitive that too little communication can impede coordination, he also demonstrates that a surplus of information can also lead to chaos. Lorenz also finds shifting regions of consensus when studying continuous opinion dynamics, the same system that is considered by the Synchronization model [43].

3.1.2 Model Description

The Synchronization model simulates the formation of a consensus opinion following a perturbation. The simulation is conducted on the 160-node trigonal network. An interesting feature of the model is that the nodes represent identical agents and that perturbations at different sites in the network have varying effects strictly because of the architecture of the relative orientation of the nodes.

Each node type has three inbound edges, with weights of 20, 30, and 50 percent. Since each node has the same inbound edges, the network can be interpreted as a tiling of identical units. This concept is illustrated in Figure 3-1.

The different node types vary in their orientation with respect to their neighbors. Each possible orientation gives a node different outbound edges. The flux is defined as the sum of the outward weightings minus the sum of inward weightings, and each node type therefore has a different flux value. The adjacency matrix for the node

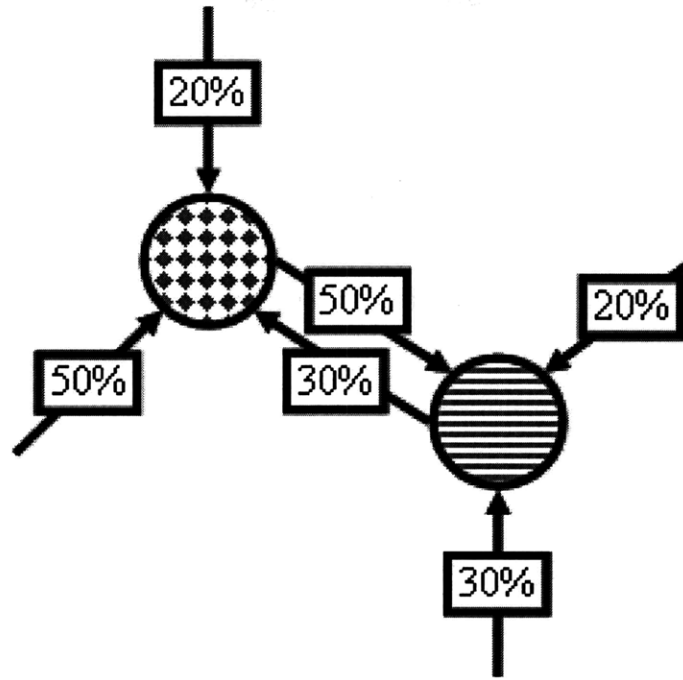


Figure 3-1: The grid is composed of identical tiles, consisting of a node and its inbound edges, oriented so as to produce varying flux scores for each node type.

types is shown in Table 3.2.

The edge weightings in the Synchronization model represent the credibility that a node associates with information inbound along that edge. Note that every node has a weighted indegree of 100% to completely account for the influence of the node's neighbors upon its opinion. The nodes therefore also share the same eigenvector centrality score. The flux values of node types A, B, C, and D are 50%, -30%, -20%, and 0%, respectively.

The network is initially in equilibrium, and each node has a baseline state value

	A	B	C	D	Flux
A	0	20%	30%	50%	50%
B	50%	0	20%	30%	-30%
C	50%	30%	0	20%	-20%
D	50%	20%	30%	0	0%

Table 3.2: An adjacency matrix for a network in which the edges represent the fractional distribution of some behavior.

of zero. A perturbation is then introduced to the network when the state value of one node spontaneously increases to 500. This event represents the appearance of an innovation in a social network or faulty data in a sensor network.

Each node in the network updates its opinion once per iteration. A node calculates the average of its neighbors' state values, weighted according to the strength of the edge leading from each neighbor to the node. The node's new state value is the mean value of this peer opinion and its own previous state value. An important implication of this algorithm is that the nodes' state values do not represent a quantity of a conserved material, since the network-wide sum of the nodes' state values will vary over time. The network eventually reaches a new equilibrium, at which every node has the same steady state value. This steady state variable is always less than the value of the original perturbation but greater than the initial equilibrium value.

3.1.3 Engineering Analysis

The Synchronization scenario models how a large community achieves consensus after one member introduces a piece of information that differs from the community's previously established shared opinion. This perturbation is invariably diminished as it diffuses through the network, but its effect is evident in the altered steady-state value of the nodes. The most influential nodes in the network are those sites at which such a perturbation will cause the greatest change in the system's final steady state, just as the most influential members of a human society are those people who have the greatest impact on the attitude of the community. The flux metric is successfully able to rank the influence of the nodes in a network according to this definition.

The relationship between the influence of a node and consensus formation during Synchronization is the opposite of the effect observed in the Conserved Flow model. The equilibrium state of the network in the Conserved Flow model does not depend on the location at which a perturbation begins. Therefore, the impact of a node in that system can be quantified by the rate at which the steady state is achieved following a shock. As described in Section 2.2.2, a more influential node is expected to distribute the unbalanced load more quickly than other nodes. However, the steady state of the

Synchronization simulation will vary with the location of the perturbation because the model does not obey a law of conservation. When a more influential node in the Synchronization model experiences a perturbation, it is able to shift the final values of the members of the network farther away from their original equilibrium values than other nodes could. Therefore, in the Synchronization model a more influential node is associated with prolonged discord rather than rapid consensus-forming because an influential node will inhibit the dampening of the original innovation or shock. The value of σ is thus greater after a perturbation at a higher-flux node, as shown in Figure 3-2.

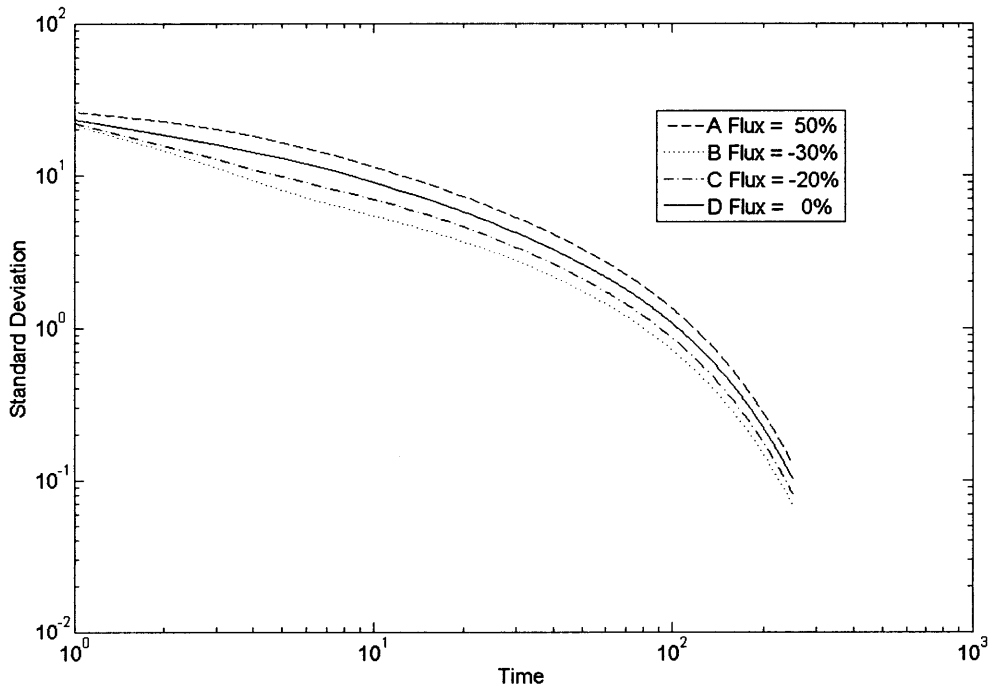


Figure 3-2: Consensus formation in a Synchronization model after a perturbation.

The mean state value of the nodes reflect the community’s opinion. The evolution of group opinion during the first 20 iterations following a perturbation at each node type is shown in Figure 3-3.

It is clear in Figure 3-3 that a perturbation at a node with a greater flux value causes the community opinion to be greater at all points in time. Since the initial state value of the nodes is zero, the opinion of the community begins at a constant

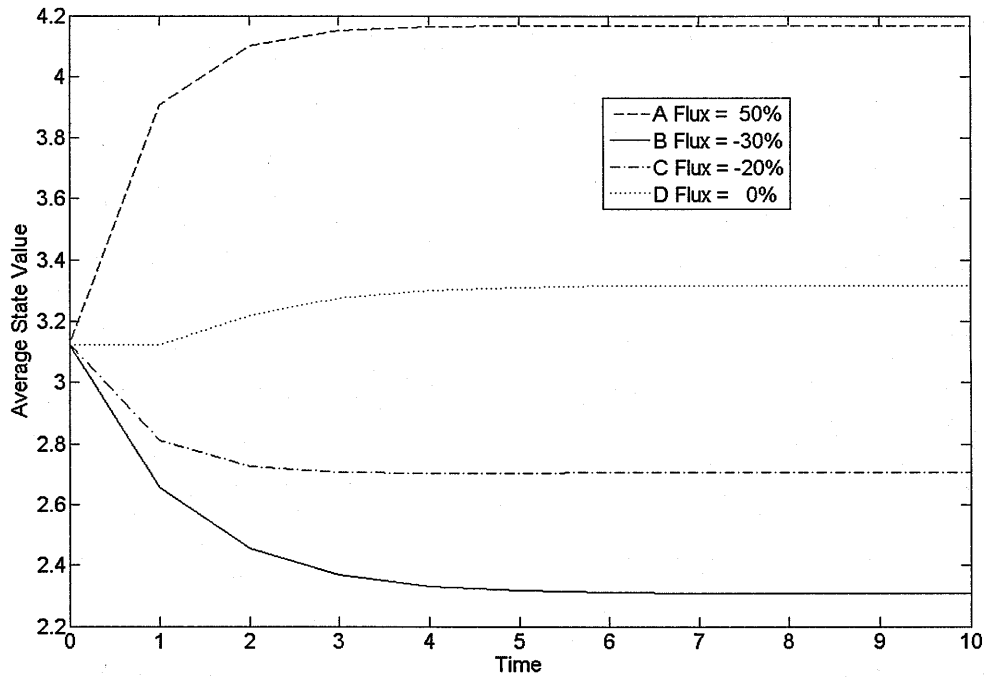


Figure 3-3: The average state value of Synchronizing nodes following a perturbation at one of four locations.

equal to the size of the perturbation divided by the number of nodes. However, it is very interesting that the average state value then decreases following a perturbation at node types B and C, which have negative flux values, and increases following a perturbation at the other sites. The declaration in Chapter 2 that nodes with negative flux values are net “followers” while those nodes with positive flux value are net “leaders” is thus upheld.

The average state value remains a constant for all future iterations past those shown in Figure 3-3. Eventually every node in the network shares this average value. Table 3.3 summarizes the correlation between each node’s flux value and the steady-state equilibrium value following a perturbation at that node type. The impact of “leader” nodes is shown by the fact that the steady-state value is greater following an innovation at a node with a greater flux value than at other nodes.

Node Type	Flux Value	Steady-State Equilibrium
A	50%	4.1667
B	-30%	2.3089
C	-20%	2.7070
D	0%	3.3174

Table 3.3: A perturbation at a node with a greater flux value leads to a greater steady-state value.

3.2 Infection

Many important systems are described by a series of discrete states. Perhaps the most prominent example is an Infection model of a population in the midst of an epidemic. Individuals are categorized as infectious, healthy, or removed [34]. There is little payoff from the tremendous increase in complexity of more nuanced descriptions such as a continuous state from initial infection to the conclusion of the disease because of the large size of the population. Despite the obvious influence of epidemics on the human condition and the long legacy of scientific inquiry into their dynamics, they are still imperfectly understood and new models are sought to better understand how they evolve and the most effective means of defending against their potentially devastating effects.

A biomimetic solution to the engineering challenge of database management is an epidemic-like algorithm. If multiple copies of a database are distributed across a network, it can be tedious to systematically update the relevant entry in each record every time the source data changes. An alternative approach is to treat the updated information as a pathogen that probabilistically infects the databases as it diffuses through the network. Previous research has indicated that this method may be more efficient than a traditional, systematic approach [19]. When implementing such a solution it is important to parameterize how quickly the update will reach a certain proportion of the network components, i.e. the rate at which the system transitions from an unstable to a stable state of consensus following a perturbation.

3.2.1 Relevant Previous Work

The formal study of epidemics may have begun when Thucydides recorded the legendary Plague of Athens in 430 BC [50]. However, the groundbreaking work in mathematical epidemiology was published in 1927 by Kermack and McKendrick [38]. The classical model derived from their work divides the community into categories depending on whether individuals are disease-free and susceptible, have the disease and are capable of infecting others, or have been removed from the population. It should be noted that “removed” in reality correlates to having recovered, died, or been placed in quarantine; the key characteristic is that these individuals can no longer infect others regardless of their own symptoms. The rates at which individuals transfer between the populations traditionally is proportionate to constant probabilities of infection and removal and the size of each population.

Kermack and McKendrick’s model assumes that the subpopulations are homogeneous and continually interacting. However, more sophisticated versions of these epidemiological models acknowledge that the members of a large community may interact with only a small subset of the population. Consequently, the instantaneous probability of a healthy person contracting the disease depends upon their social context. In cases of such limited social interaction it is intuitive to represent the transmission of the disease among individuals as a network-based phenomenon [76]. The nodes of the network may also represent families or other social organizations within which the rate of disease transmission is orders of magnitude greater than it is between such groups.

Network analysis techniques are particularly appropriate when modeling the spread of sexually transmitted diseases. It is possible to explicitly map and analyze a network of sexual relationships, in contrast to the much more anonymous means by which most other diseases are transmitted. Rothenberg and Narramore, for example, have studied records of syphilis patients in Tennessee to identify relatively isolated subpopulations of patients between which a small number of individuals act as critical bridges. Their objective is to better allocate health care personnel to each population

[55]. Computer simulations of gonorrhea transmission in a population capitalize on network analysis techniques to determine necessary factors for the establishment of a perpetual infection and the amount of sampling needed to gauge the spread of the disease [28]. Network analysis techniques have been especially popular in studies of human immunodeficiency virus (HIV). These methods have helped scientists to recognize the characteristics of acquired immune deficiency syndrome (AIDS) during the original discovery of the disease [26].

The transmission networks of some non-sexual diseases may also be mapped explicitly because of the small scale of the outbreak. The severe acute respiratory syndrome (SARS) outbreak that began in 2002 is one such example. It has been observed that despite the high population density of the affected cities, such as Singapore and Hong Kong, a small number of so-called “superspreaders” that represent at most 3% of the infected population have been responsible for 80% of new infections. Masuda and Konno have investigated the characteristics of a superspreader and what properties of the disease lead to this type of transmission network [47]. Their studies highlight the need for measures of influence in transmission networks that do not rely on a thorough understanding of the nature of the disease.

Scientists have investigated which centrality metric is most useful for epidemiological research. Rothenberg et al. used eight different measurements to determine the centrality of individuals at high-risk for HIV infection in one city [56]. Rothenberg et al. noted that the actual value of centrality scores are likely to be very noisy and therefore the data should be used only to rank the importance of particular nodes rather than compare the relative magnitude of their influence. They also observe that weighted networks are necessary to incorporate the varying levels of risk associated with a relationship between two individuals according to their shared activities. Despite this limitation, Rothenberg et al. have found that the weighted and non-weighted metrics produce very similar results. However, in a separate analysis of a different community of high-risk individuals by Bell et al., weighted measures are markedly more accurate than unweighted techniques [2].

Discrete-state network models are also relevant to non-medical disciplines. These

simulations are mathematically similar to epidemiological networks. For example, Sznajd-Weron and her colleagues have conducted a series of network studies of the Ising spin model, which describes the binary state of electrons [69][68]. The purpose of these inquiries is to determine how the value of an electron will affect its neighbors. Issues of centrality and influence are therefore at the core of their investigation. Intriguingly, a key parameter that they consider is the “outflow” of a node [63]. Recall that the weighted outdegree of a node determines its flux value if the weighted indegree, and hence eigenvector centrality, of all nodes are identical. Sznajd-Weron’s method is separate and distinct from the flux metric, but it is nonetheless intriguing that the latest studies of influence in networks also concentrate on the outward edges of a node when inward links traditionally have been the focus of such research.

A team at the Xerox Corporation working to update multiple copies of a database located at the nodes of a network have transferred discrete-state models to engineering design [19]. The team, led by Demers, has discovered that the information at the nodes may be most reliably updated by a disease-like algorithm if no central list exists of the nodes to send an update to. The model used by Demers et al. that is most relevant to this project is “rumor mongering”. All nodes are initially “ignorant”, analogous to a healthy, susceptible population. After a site receives an update, it randomly selects a neighbor to attempt to pass the “hot rumor” to. If after a certain number of iterations the infected site continuously finds that its neighbors already knew the update, then the update becomes “old news” and the node no longer attempts to pass the update along. This process is similar to the recovery of infected nodes in the disease model.

3.2.2 Model Description

An Infection model demonstrates the insight that flux metric could offer to studies of epidemic networks. The simulation executes on the same 160-node trigonal network of identical tiles as the Synchronization model. In the Infection model, inbound edge weightings in Table 3.2 represent the amount of time that an individual is potentially exposed to an infection due to a particular relationship. It is assumed that direct

contact is not necessary for infection, but rather that the pathogen is associated with an environment shared with a neighbor. For example, consider a married couple in which one spouse spends only 20% of their day in their home but the other spouse spends 60% of their time there. Recall that because the node types shown in Table 3.2 all have the same weighted indegree and eigenvector centrality score. This particular graph and adjacency matrix are admittedly abstract, and are meant only to investigate the possibility of using the flux metric to gain insight into studies of epidemics rather than to analyze a particular real-world scenario.

The nodes are always either susceptible, infectious, or removed. Initially, only one node is infectious and the remainder are susceptible. The behavior of a node at each iteration depends on its state at that time. A susceptible node may be infected by any of its infectious neighbors. The portion of time spent by an individual in each environment shared with an infectious node is multiplied by the constant rate of infection per unit time to determine the probability that the node will be infected that iteration, and a random number generator determines if this event occurs. There is a constant probability that an infected node will be removed. A removed node does not change its state nor alter the state of its neighbors, but remains in the network. Therefore, the network structure remains constant even though the state of each node changes. A node's neighbors are predetermined and constant. This modification is a simplification of the epidemic-like algorithm used to update simultaneous databases, in which a node may randomly select any other connected node to infect. The change is necessary for analytical purposes to ensure that the nodes each have a constant flux value throughout the experiment

3.2.3 Engineering Analysis

The Infection model is probabilistic, a feature unique among the simulations presented in this project. Adjusting either the probability of infection or removal will change the simulation dynamics. Let α be the probability of removal and β represent the probability of infection. Both α and β may be any real number equal to or greater than zero and less than or equal to one. Note, however, that an infection rate of zero

is uninteresting because it prohibits the disease from spreading beyond the initial perturbation.

A wide-spread outbreak would be expected to occur more frequently following the spontaneous infection of an influential node than another node. The scale of the outbreak at any point in time is measured by the size of the infectious population. Note that σ is not relevant to the Infection model due to the small number of discrete state values permitted. Figure 3-4 shows the change in the size of the infectious population after an outbreak begins at each node type when α is 0.05 and β equals 0.40. Each plot shown is the result of a single probabilistic simulation.

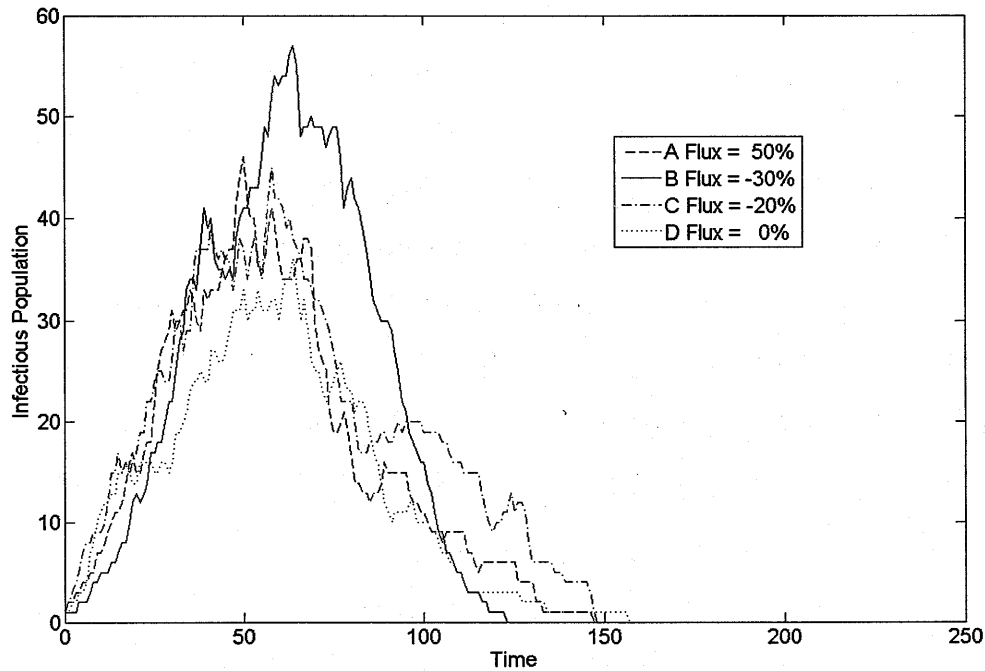


Figure 3-4: The size of the infectious population following a spontaneous outbreak at four node types in the Infection model.

As can be seen in Figure 3-4, there is no clear correlation between the flux value of the initially infected site and the ensuing outbreak. The flux value of the node at which the outbreak begins does not appear to be an accurate indicator of the size of the infectious population at a given time. It is possible that the flux value of a node could have a more pronounced impact if the values of α and β were different.

Any consequence that flux has would be most obvious in a comparison of node type A, which has the greatest flux value at 50%, and node type B, which has the least flux value at -30%. Suppose that the maximum size of the infectious population after node type X is infected is designated I_X and the time at which this peak value occurs is T_X . The difference in the history of the infectious populations after an outbreak begins at each node type can then be quantified by the distance between these maxima,

$$distance = \sqrt{(I_A - I_B)^2 + (T_A - T_B)^2} \quad (3.1)$$

Figure 3-5 displays the effect of α and β on the distance between the maxima. To increase the statistical significance of the results, the location of the maxima used to calculate the distances shown in Figure 3-5 are actually the average values observed after ten identical simulations.

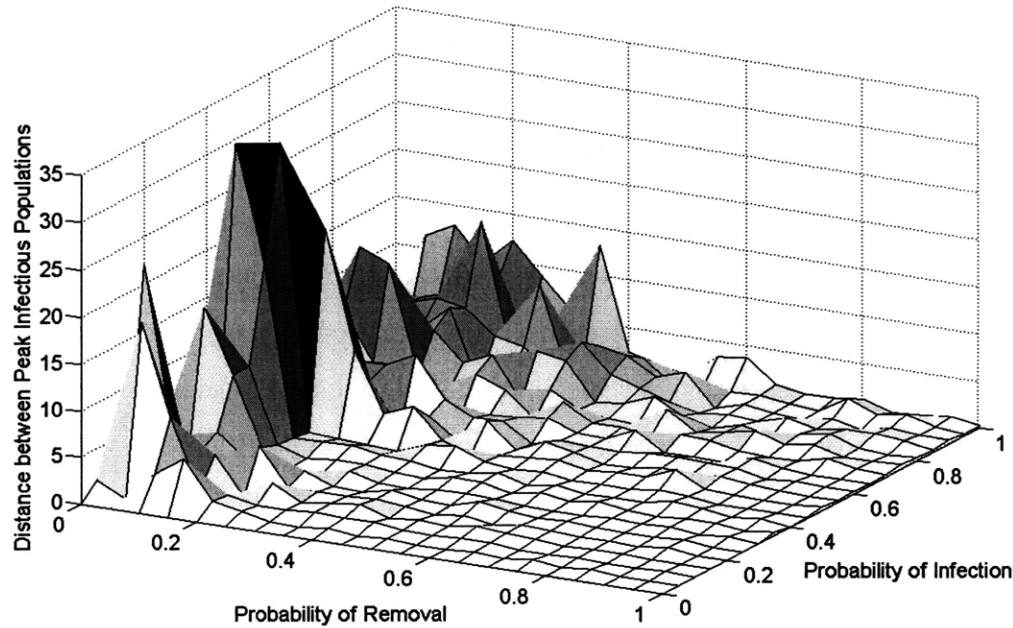


Figure 3-5: The effect of the flux value of the initial site of an epidemic as a function of the infection and removal rates.

Further investigation reveals that the flat regions in Figure 3-5, where the effects of an outbreak that begins at node types A and B are identical, are conditions in

which no epidemic actually occurs. The initial node recovers before infecting any of its neighbors. Kermack and McKendrick's differential equation model of disease transmission also predicts threshold conditions that determine whether an epidemic will grow or dissipate. Their conclusion is that the initial rate of growth of the population of infectious individuals will only be positive if the ratio of the probability of removal to the probability of infection is greater than the initial size of the susceptible population.

Kermack and McKendrick's threshold rule must be modified for application to a network because the total population is not free to mix. It is logical to assume that the initial size of the relevant susceptible population would only be three because the initially infected node is in contact with its three neighbors. It can be seen in Figure 3-5 that the region in which $\alpha > 3\beta$ is perfectly flat and no true epidemic follows the appearance of an infectious node. Therefore, Kermack and McKendrick's threshold rule for the development of an epidemic holds true for the Infection model.

Unfortunately, Figure 3-5 reveals no other relationship between α , β , and the impact of the flux value of the epidemic's origin. Even in regions that are associated with much volatility, such as where the probability of infection lies between zero and 0.20, further investigation reveals that the relationship between epidemics that originate at node type A and those that begin at node type B is not consistent. For example, the results shown in Figure 3-4 were for an α of 0.05 and a β of 0.40.

The difficulty in gaging the effectiveness of the flux metric to predict the dynamics of an epidemic is that models of disease propagation are probabilistic. The statistical significance of any given set of results is somewhat unclear, and computing resource limitations prevent the execution of very large numbers of simulations. Whatever effect the flux value of the initial site of an outbreak might have on the ensuing epidemic is nullified by probabilistic effects.

3.3 Resource Sharing

Many intriguing network phenomena involve Resource Sharing within a community. In these systems a physical good moves among nodes with identical structures. The Resource Sharing model therefore obeys the same conservation laws as Conserved Flow but is composed of nodes whose varying characteristics emerge as a result of their orientation to their neighbors, an architecture similar to that of the Infection and Synchronization models. The Resource Sharing model represents actual systems from areas as diverse as robotics, finance, and decentralized computing. In engineering, the importance of this structure is that the dispersal of a material through the network can be engineered by altering the network's architecture without the need to create custom units with varying connection weightings nor the need for some units to have more connections than their peers.

The details of a particular implementation of the general Resource Sharing model would be very intricate and specific to that scenario. However, the basic concept of the system are explained by a simple allegory. Suppose that there is a community of children, one of whom spontaneously receives a large quantity of candy. The child wants to enjoy the candy, but also knows that he should share with his friends to maintain his relationships with them. The child therefore keeps a certain percentage of the candy for himself and doles out the remainder to his neighbors according to the value that he places on that friendship. These secondary recipients in turn keep a percentage of the candy and pass the remainder along to their friends, including the original recipient, according to their social priorities. Since the network is fully connected, eventually a steady state will be reached and the amount of candy that each child has will not change despite the incessant flow of candy among them. The path to this equilibrium depends on the influence of the child that received the original gift.

	A	B	C	D	Flux	Eigenvector Centrality
A	0	50%	50%	50%	-50%	0.1029
B	20%	0	30%	20%	30%	0.0570
C	30%	20%	0	30%	20%	0.0668
D	50%	30%	20%	0	0%	0.0819

Table 3.4: An adjacency matrix representing the fractional distribution of a conserved good.

3.3.1 Model Description

The Resource Sharing model uses a network of 160 identical nodes on a trigonal grid, similar to the structure of the Synchronization and Infection models. However, the nodes of the Resource Sharing model have the same outbound edges rather than identical inward edges. The nodes therefore may have varying eigenvector centrality scores. Even if an existing centrality measure can rank the relative significance of each node, this scenario is still pertinent to this study because the applicability of the flux metric to such a ubiquitous situation must be verified before advocating the widespread of the metric.

Table 3.4 contains the adjacency matrix describing the relationship between each of the node types. Table 3.4 is the inverse of the adjacency matrix shown in Table 3.2 that is used for the Infection and Synchronization models. This relationship emphasizes that, in contrast to the structure of the other models, the weighted outdegrees rather than the weighted indegrees of the Resource Sharing nodes are identical. However, there is no intrinsic limitation on the edge weightings other than that each node has a weighted outdegree of 100%.

Table 3.4 also shows the eigenvector centrality score of each node type when it is located in a 160-node trigonal network. Note that the ranking of the nodes according to their eigenvector centrality is identical to the flux score ranking. However, the node type with the greatest eigenvector centrality score and weighted indegree has the least flux score. This contrast occurs because every node has the same outdegree and each node's flux is calculated by subtracting its indegree from this constant.

The flux values of node types A, B, C, and D are respectively -50%, 30%, 20%,

and 0%. During each iteration of the simulation, a node retains a set percentage, κ , of the material in its possession. The node then distributes the remaining material to its neighbors according to the relationships in Table 3.4. Each node is modeled as being equally selfish, i.e. the value of κ is the same for each node type.

3.3.2 Engineering Analysis

In this model a conserved good is relayed among the nodes. Therefore, σ is an appropriate measure of the influence of a node per the discussion in Section 3.1.3. The network-wide standard deviation for 250 iterations following an initial perturbation of 250 units at each of the four node types is shown in Figure 3-6. In this scenario κ is set to 25%.

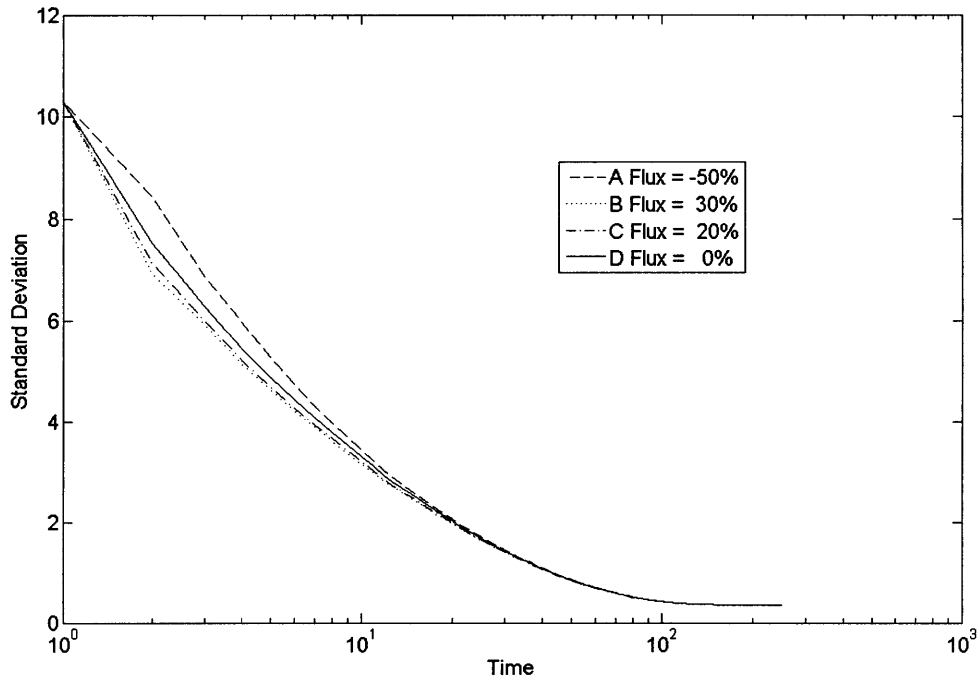


Figure 3-6: The stabilization of a Resource Sharing network after one node receives a perturbation of 250 units.

As can be seen in Figure 3-6, equilibrium is achieved slightly more quickly after a perturbation at a node with a greater flux value than at a node with a lesser flux value. Each node reaches the same steady-state value regardless of where the perturbation

Node Type	Flux Value	Steady-State Value
A	-50%	2.0833
B	30%	1.1545
C	20%	1.3535
D	0%	1.6587

Table 3.5: The Shared Resource accumulates at nodes with the least flux values.

begins. Interestingly, not only is the steady-state value of a node dependent on its flux type but the ranking of the equilibrium values also corresponds to the flux ranking of the nodes. Therefore the node with whom relationships are most valued eventually and hence has the least flux score receives the greatest amount of the resource. The flux values and final state values for each node type are summarized in Table 3.5.

A key parameter in the Resource Sharing model is κ , the percentage of the material that each node keeps for itself. This parameter can be tuned to elicit different system behaviors on a predefined network. For example, suppose that an engineer is interested in γ , the difference in σ at a given point in time if the network has been perturbed at a node with the greatest or the least flux value. The behavior of γ is therefore the maximum effect that the location of a perturbation can have on its outcome.

Note that in this scenario γ is the difference in σ following a perturbation at node types A and B from Table 3.4. The significance of γ is relative to the magnitude of the σ values at that time. The maximum value of the ratio of γ to σ_A throughout time for each possible value of κ is shown in Figure 3-7.

The time at which the maximum value of γ occurs also varies with κ . Results not shown here indicate that these properties do not vary with the size of the perturbation, only the network architecture and the edge weighting between nodes. The abrupt change in behavior observed in Figure 3-7 when κ equals 35% is an example of the complex behavior that can arise from even simple network dynamics. It is an especially interesting phenomenon because no similar bimodal behavior can be observed in a plot of the time at which the maximum gain in stabilization is achieved, which is shown in Figure 3-8.

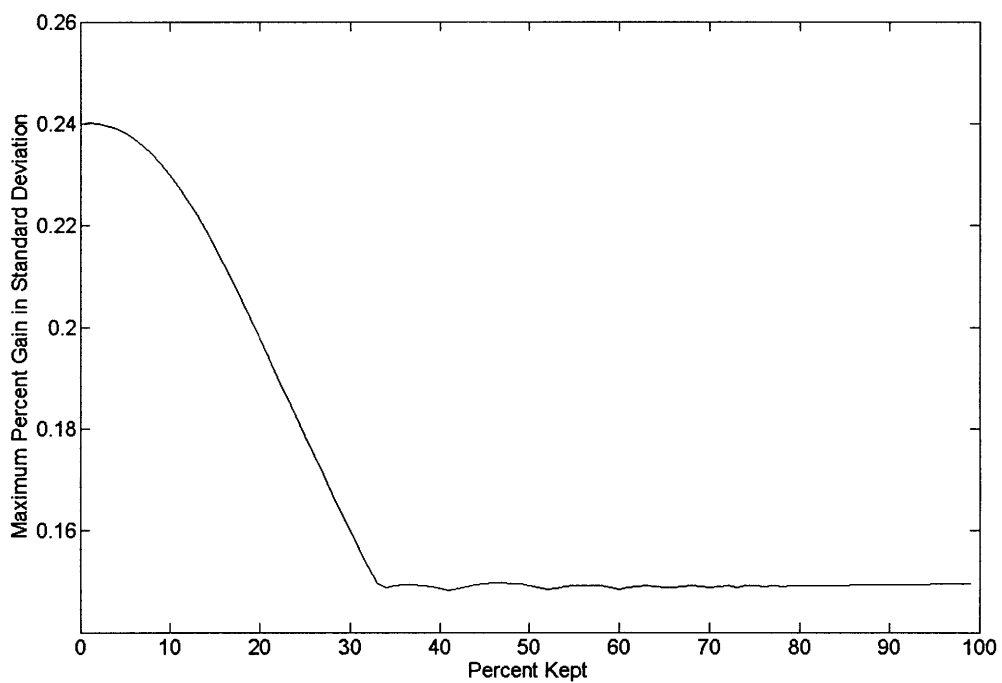


Figure 3-7: The maximum decrease in σ achieved in the Resource Sharing model following a perturbation at a node with a flux value of 30% versus a node with a flux value of -50%.

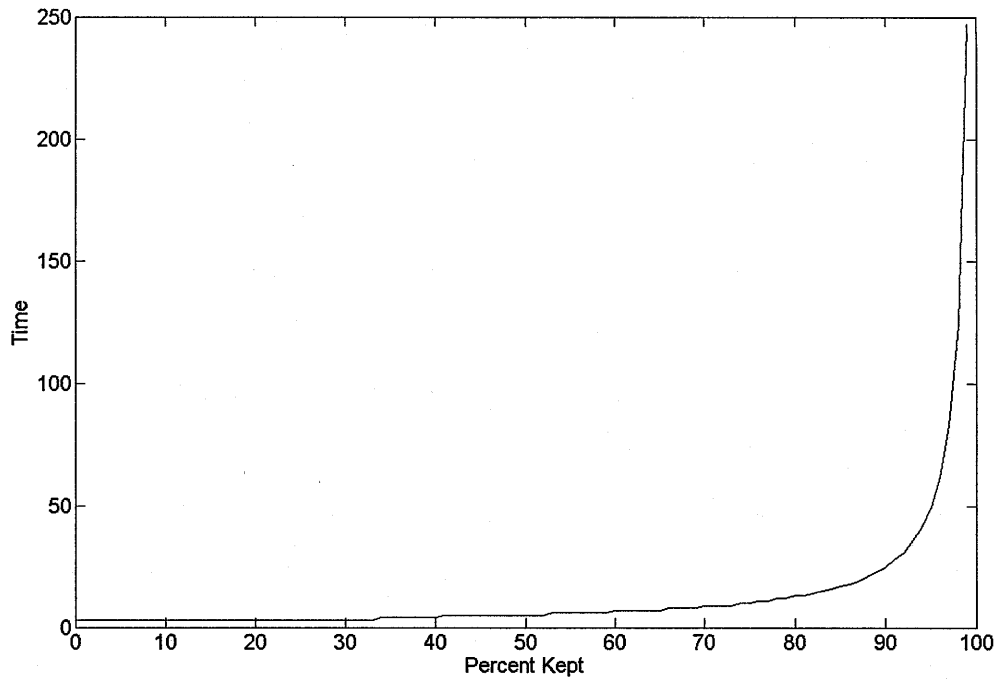


Figure 3-8: The time at which the ratio of γ to σ_A is the greatest.

An engineer may want to examine the gain in consensus formation throughout time following a perturbation at the higher-flux node in comparison the the effect of a perturbation at the lower-flux node. This value is the percent difference in the integrals of σ_A and σ_B , and is shown as a function of κ in Figure 3-9.

Figure 3-9 shows that the effect of a perturbation's origin in the Resource Sharing model is more pronounced the greater the value of κ is. This is another insight that the flux metric and the Resource Sharing model offer into the design of a potential system.

Conclusion

The flux metric has been shown to be a useful tool for analyzing the scenarios represented by the Conserved Flow, Synchronization, Infection, and Resource Sharing models. The flux metric is thus a valid method of ranking the influence of nodes under a large variety of network dynamics with practical applications to real-world engineering and social challenges. While the flux metric has noted shortcomings,

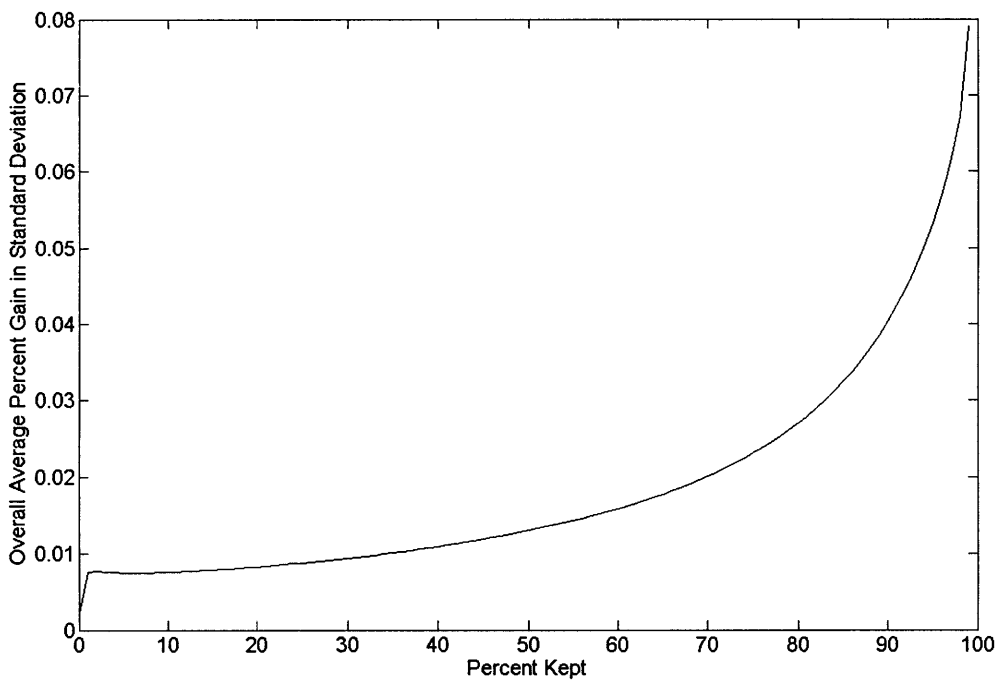


Figure 3-9: The overall improvement in network-wide standard deviation following a perturbation at a node with a flux value of 30% versus a perturbation at a node with a flux value of -50%.

such as its inability to be very useful in probabilistic settings, it has the noted advantage of differentiating among nodes with constant weighted indegree and eigenvector centrality scores.

Chapter 4

Engineering Design through Network Analysis

Introduction

Networks relationships need not be limited to the passing of an object or opinion among nodes. Rather, almost any quantified relationship among a system of nodes can be formalized as a mathematical graph. A novel challenge to which network analysis techniques could be fruitfully applied is the design of complex systems. The concept of system design as a network is evidently unique; the most similar published work is Liu et al.'s very recent application of network techniques to data envelopment analysis [42]. Their approach is to express the relationships between data management units as a network and to then use the eigenvector centrality score of each unit to determine which is the most efficient. Data envelope analysis is an operations research method that relates the inputs and outputs of a system. This technique is similar to engineering design, which seeks to relate the objectives of a design to the subcomponents that comprise the system. The relationship between the missions and subsystems of a product can be mapped as a graph. Measures of nodal influence, particularly the flux metric, can then be used to prioritize design efforts or contribute to cost studies. In this chapter these techniques are demonstrated by a case study, the design of autonomous underwater vehicles (AUVs) for the offshore oil industry. This

study fulfills the goals of demonstrating the use of networks for design and verifying the application of the flux metric to non-conventional networks.

4.1 The Design Problem as a Complex Network

Complex systems are composed of many subsystems. It is very important to understand which subsystems are more important than others when designing against failure, prioritizing supply streams, allocating budgets, and scheduling research and development. However, ranking the significance of each subsystem is very difficult because only the relationships between coupled subsystems are known directly. This problem is often exacerbated in a corporate setting by the presence of subcontractors, reducing the level of communication between the design teams assigned to the various subsystems.

The key to understanding the relative significance of each subsystem is to recognize that the subsystems are related via the system objectives that they fulfill. It is much simpler to determine the direct relationship between a subsystem and a mission than the more abstract connection between two subsystems. Through the techniques and theories of network analysis, this knowledge may then be used to better understand how the system as a whole.

Recall from Section 1.1.1 that in a bipartite network relationships exist between two classes of nodes, but that no edges connect the nodes of a single class. The special significance of weighted, directed networks is due to their appearance during the decomposition of a bipartite graph to a unipartite graph as shown in Figure 1-2. Bipartite graphs are a natural representation of many complex systems because they divide a large system of nodes into smaller units that may be better understood. For example, large-scale social systems are often comprised of smaller, overlapping organizations. Since many members may belong to more than one organization, it can be difficult to determine their relative influence on the entire system even if their role within each individual organization is well established. However, modeling these local measures of importance as the edges linking the individuals to their organizations in

a bipartite graph, decomposing this graph to its unipartite form, and applying the flux metric would reveal the system-wide importance of each member.

A bipartite graph is therefore an appropriate representation of a complex system in which the correlation between each subsystem and one or more missions are well-known, but the relationships among subsystems and missions are more abstract. This project focuses on one particular and practical interpretation of weights. The weighting of the edge connecting a particular mission and subsystem corresponds to the probability that the subsystem will reach peak capacity during that mission. This peak behavior infers that the mission could achieve more if it were not limited by the subsystem. If a subsystem is not used in a mission, then the probability of peak performance occurring is zero and no relationship exists between the mission and that subsystem.

All devices have some sort of quantified design limit, whether it is the maximum thrust exerted by a propulsor or the peak pressure that a vessel can withstand. Although a system can be described by many different characteristics, the appropriate parameter to consider for this approach is the device trait that limits a subsystem's contribution to the mission. For example, if the current mission is taking sports photographs then shutter speed may be the limiting trait of a camera. However, if the objective is to photograph dangerous animals then the power of the telephoto lens would be more appropriate.

Note that the probability of any particular mission type occurring is known and that since all possible missions are identified, these probabilities sum to one. The weightings of the edges between subsystems and missions in the bipartite graph are the probability of the subsystem peaking, given that the relevant mission is underway, multiplied by the probability of that mission occurring. Hence, the weighting is the overall probability of that subcomponent peaking during that mission.

For example, suppose that P_Y is the probability that mission Y will occur and X_Y is the probability that subcomponent X will peak during mission Y . Furthermore, there are two missions α and β and three subcomponents A , B , and C that are utilized during every mission. The bipartite graph depicting this scenario is shown in

Figure 4-1. Only the weighting of one edge is labeled for clarity.

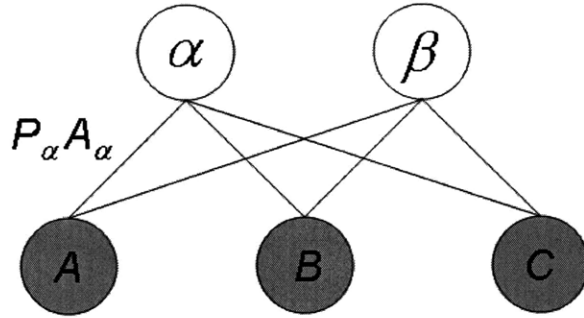


Figure 4-1: A bipartite graph modeling the relationships between the subsystems A , B , and C and objectives α and β of a design.

4.1.1 Ranking Subsystems

After the bipartite network describing the system is defined, it can be decomposed to form a unipartite graph representing the relationships among subsystems. Recall from Section 1.1.1 that each edge between a subsystem and a mission in the bipartite graph is expanded to become a set of edges leading from that subsystem to every other subsystem that had been originally linked to the mission. The network of subsystems produced can be analyzed to determine the relative influence of each subsystem.

The strict definition of the edge weightings in the design network permit a formal interpretation of the flux metric as a measure of the influence of each node. Suppose that a product utilizes three subsystems to achieve two different missions, as described in Section 4.1. The flux of subcomponent A is defined as

$$flux(A) = outdegree - indegree \quad (4.1)$$

$$flux(A) = (2P_\alpha A_\alpha + 2P_\beta A_\beta) - (P_\alpha B_\alpha + P_\beta B_\beta + P_\alpha C_\alpha + P_\beta C_\beta) \quad (4.2)$$

$$flux(A) = P_\alpha(2A_\alpha - B_\alpha - C_\alpha) + P_\beta(2A_\beta - B_\beta - C_\beta) \quad (4.3)$$

The flux of a subcomponent is a measure of the node's influence because it indicates how much more likely the subsystem is to peak than the other subcomponents involved in the same mission. Nodes with positive flux scores are more likely to peak

than the other subsystems utilized in the same missions, while nodes with negative scores are less prone to peaking than other subsystems. Peaking is normally undesirable because it infers that the design limits of the subsystem add constraints to the potential achievements of a mission. Furthermore, the wear and tear on a device dramatically increases when the subcomponent is operating at its maximum possible performance level.

A sudden change in a subsystem's performance is a perturbation analogous to the network shocks discussed in Chapters 2 and 3. The overall capabilities of the system will be affected most by the capacity of an influential subsystem as measured by the flux metric. A subsystem will tend to have a large flux value if it is very likely to peak in any missions that depend on it, even if the subsystem relates to fewer missions than other subcomponents. Allocating resources towards expanding the maximum possible performance of such a subcomponent would significantly expand the range of tasks that could be accomplished during the missions.

4.1.2 Ranking Missions

The bipartite graph depicting the missions and subsystems may also be decomposed into a unipartite network of missions. The mathematics of this procedure are identical to those used to form the unipartite network of subsystems, except that the flow of the edges is reversed. When ranking missions the unipartite edges travel from a mission to all other missions that are connected to a shared subsystem in the bipartite network. Considering the same example from Section 4.1, the flux of mission α is thus

$$flux(\alpha) = outdegree - indegree \quad (4.4)$$

$$flux(\alpha) = (P_\alpha A_\alpha + P_\alpha B_\alpha + P_\alpha C_\alpha) - (P_\beta A_\beta + P_\beta B_\beta + P_\beta C_\beta) \quad (4.5)$$

$$flux(\alpha) = P_\alpha(A_\alpha + B_\alpha + C_\alpha) - P_\beta(A_\beta + B_\beta + C_\beta) \quad (4.6)$$

A mission with a positive flux score more frequently causes subsystems to peak than do other missions that share the same subsystems. Peak-capacity behavior tends to be correlated with fatigue and other damage, and therefore is indicative of

maintenance costs. However, when repairs are conducted all missions that share the relevant subsystem benefit. Missions with positive flux values are partially responsible for the operating costs of other missions while missions with negative flux scores benefit from repairs directly associated with other missions. This insight is applicable when choosing how much to charge customers if bills are itemized per mission or checking that the profit associated with each mission type corresponds to its cost. Since subsystems are shared between missions, garnering this financial insight without a network-based approach would be extremely difficult.

4.2 Case Study: AUVs for the Offshore Industry

The application of the flux metric and other network techniques to a current engineering design topic will be demonstrated by an analysis of the use of autonomous underwater vehicles (AUVs) in the offshore oil industry. AUVs are evolving rapidly from experimental concepts to commercial products. AUVs offer solutions for maritime missions too deep, dangerous, or extensive in space or time for human divers and at a lower cost than other robotic alternatives. The offshore industry in particular abounds with opportunities where the astute application of AUVs to perennial challenges would increase both cost savings and human safety. However, AUVs are truly complex systems whose various subsystems may interact to accomplish their objectives while simultaneously competing for the scarce resources available aboard the vehicle. The nonlinear relationships and indirect causalities that thus arise while developing the vehicle's subsystems encourage the use of rigorous quantitative techniques, such as this network-based approach, to augment the traditionally qualitative design process. This case study includes a description of the AUV scenario, quantitative results of a survey given to members of the oil and gas community, and the interpretation of the data as flux values with a comparison to another ranking scheme.

4.2.1 Overview of Deep-Sea Technology

A thorough accounting of the needs of the offshore industry and the current capabilities of AUVs is necessary to explain the scientific and industrial basis of the subsequent network-based study. The deep ocean offers great promise for discovery and commercial opportunity, but also presents significant challenges that distinguish this environment from other regions of exploration.

The density of sea water causes crushing pressures to quickly increase with depth. For every ten meters of depth, the water pressure increases by approximately one atmosphere [52]. The human body cannot tolerate more than four atmospheres of pressure, and even such limited depths mandate careful attention to depressurization rates to limit decompression sickness [48]. The challenge posed by the increasing pressure escalates at greater depths, because conventional solid-state electronic components begin to fail at high pressure. This phenomenon is the one of the most significant reasons for the high cost of deep-sea exploration. Simple systems that would cost hundreds of dollars to operate on the surface escalate in cost to tens of thousands of dollars in order to be functional thousands of meters below the sea.

The single greatest challenge to developing deep-sea systems is posed by the medium itself. Water attenuates electromagnetic radiation; therefore radio communication is unreliable at even moderate depths. Technologies developed to explore other environments, such as space probes, unmanned aerial vehicles, and all manner of terrestrial craft, are therefore rendered useless in the ocean. The Global Positioning System (GPS), which has become a mainstay of computer-assisted navigation, is completely unavailable underwater. While acoustic solutions have been developed to permit limited communication between surface craft and deep-sea vehicles, such systems lack the bandwidth and range to permit the persistent communication assumed possible in other realms. Deep-diving underwater vehicles are classified according to the degree of involvement of a human pilot during their operation. At one extreme, a deep submersible vehicle (DSV) actually has a human crew onboard. The crew typically consists of a pilot and one or two scientific observers. Although a DSV

is untethered and can navigate under its own power, the limitations of life-support systems and operator endurance restrict the horizontal range of the vehicle. A DSV therefore is typically deployed by a surface-based support vessel directly above the region of interest, to which it descends and returns in approximately eight hours. The main advantage of a DSV over a robotic system is the increase in situational awareness and mission flexibility afforded by the presence of a human pilot onboard. However, the cost associated with maintaining the life-support systems aboard a DSV and mitigating the significant safety risks incurred by each expedition to the deep make DSVs much more costly than other exploration vehicles [14].

A remotely operated vehicle (ROV) is in principle very similar to a DSV, except that the pilot has been removed from the confines of the vehicle and teleoperates it from the support ship. The pilot therefore has approximately as much control over vehicle manipulators and thrusters as he or she would aboard a DSV, since such mechanisms are electronically coupled to human input devices in either vehicle, their situational awareness of the vehicle's operating conditions and physical surroundings is limited to video and sensor readings. The US Navy developed the first true ROV, the Cable-Controlled Undersea Recovery Vehicle (CURV). Although this vehicle only dived to a moderate 600 m, in 1966 it famously retrieved a nuclear weapon lost off the coast of Spain [12]. By the 1970s, ROV development was evenly divided between commercial and defense applications and by the end of the decade few DSVs remained [14].

Signals travel along a tether linking the ROV directly to the support ship. While the umbilical limits the range and maneuverability of a ROV, it also serves as a conduit of energy from the ship's power plant. While these resources theoretically grant a ROV unlimited operating time, the constant supervision required by the surface-based pilot in practice limit the utilization of the vehicle [14].

AUVs replace the need for continuous human oversight with artificial intelligence. This permits the vehicles to operate around the clock and eliminates the need for a tether. The sophistication of the onboard programming can vary anywhere from blindly executing a predefined series of navigational maneuvers to dynamic responses

to changing conditions underwater [16]. Initial AUV development was driven by military needs, similarly to the early history of ROVs. The first deep-sea AUV, Epaulard, was built in France during the 1970s and dove to 6,000 m. In the 1980s, the US Navy pursued AUV technology in order to meet the demands of anti-submarine warfare missions [14].

While the absence of an umbilical imposes power limitations on an AUV, this feature also permits access to regions too constricted for ROVs and too dangerous for DSVs. Furthermore, advanced power schemes may allow an AUV to automatically refuel at either mobile or permanently moored solar-powered docking stations. Singh et al. have demonstrated that an AUV is capable of homing to a docking station and then communication by satellite with a scientific base station [62]. Since AUVs do not require a support ship, they may explore regions inaccessible to surface craft such as the water beneath the ice caps. The large capital investment but low operating costs of AUVs also make them ideal for repetitive, routine missions, such as those encountered in the offshore industry [14].

4.2.2 AUV Subsystems

An AUV is the compilation of many individual subcomponents integrated together to serve a specific mission. Each AUV is custom built to meet its customer's specifications, and even major AUV companies such as Bluefin Robotics still report sales in the single digits [13]. Because of the young age of the AUV industry, many of the devices found aboard an AUV are not primarily marketed as AUV components. While this tendency means that engineers must be diligent to ensure that new parts will function appropriately, it also indicates that there are many opportunities for innovation and development in the AUV industry. Virtually any component conceivable can be integrated into an AUV, so long as it can function underwater and at high pressure, making the possibilities of future AUV design very promising. In order to categorically analyze potential AUV subcomponents, they are grouped according to their function aboard the vehicle, from pressure-resistant materials to manipulators and other invasive tools.

Materials

AUVs require materials that can withstand the pressure changes, temperature variations, and corrosive effects of transiting the water column. Both the cost of materials themselves and the difficulty of manufacturing with the stronger materials cause the cost of materials to increase with the maximum depth of a vehicle. Metal compounds with a high yield stress, such as high-strength steel, aluminum alloys, and titanium alloys, are susceptible to stress corrosion cracking. While cathodic protection can guard against such corrosion, this technique makes all metals except aluminum alloys susceptible to hydrogen embrittlement [46]. Furthermore, metals are typically denser than alternative materials, which is undesirable because vehicles are designed to be as light as possible to conserve onboard power supplies and reduce the need for deep-rated flotation, which is quite expensive. However, metal compounds are still incorporated into some AUV designs. For example, enhanced methods using titanium alloys have been developed in Russia and Ukraine despite the six-fold cost increase associated with machining and welding titanium in comparison to aluminum [14]. Non-metallic alternatives include carbon fiber reinforced composite (CFRP) and ceramic hull structures. A main disadvantage of CFRP is its anisotropy, requiring an engineer to be very cognizant of the orientation of the fibers relative to the load in a given component. While ceramic hulls have been under development since the 1960s, they are still associated with high manufacturing costs. Furthermore, ceramic materials are very brittle so localized stresses as unremarkable as those present at an O-ring groove must be avoided [64]. Acrylic plastics are used for viewports rather than glass because acrylic is easier to machine, has more reproducible physical properties, and has a more predictable point of catastrophic failure. Windows are generally conical, rather than the easier to manufacture flat shape, because the conical shape adds greater strength. Acrylic is strongest when under compression, so primary structures made of acrylic are always spherical to ensure uniform hydrostatic loading [46]. Glass spheres may also be used when lives are not at risk. For example, MIT Sea Grant AUV Lab's Odyssey IV has two 17" glass spheres at its core that house its

electronic systems to depths of 6,000 m [40]. A wide variety of materials are also available for controlling the buoyancy of an AUV. An AUV is trimmed to be slightly positively buoyant so that the vehicle can be recovered following propulsion system failure. Since the density of sea water changes with depth, there is a degree of cost effectiveness associated with designing an AUV to operate primarily at one depth [64].

Propulsion

The majority of AUVs are driven by propeller assemblies [60]. Rather than rely on a rudder or other control surfaces, an AUV frequently has multiple thrusters oriented to provide force in each of the primary directions of motion for that particular AUV [27]. Some AUVs also have azimuth thrusters that can independently rotate [37]. DSVs and ROVs use nozzles to enhance propeller efficiency, a trend not seen on early AUVs because their propellers were not powerful enough for such hydrodynamic effects to become significant [14]. However, the latest AUVs generate more thrust than their predecessors and consequently are equipped with such nozzles.

In 1996, the National Research Council deemed AUV propulsion a mature technology that was unlikely to yield anything but marginal improvement. However, the past decade has heralded several creative breakthroughs that capitalize on the advantages of AUVs' diminutive size in relation to most other maritime vehicles. The first of these innovations was the buoyancy-driven glider, an extremely energy-efficient vehicle. Traveling at the relatively slow speed of 0.5 knots for up to one year, the winged AUV automatically changes its buoyancy via a series of bilge pumps. As it rises and falls in the ocean the wings of the AUV convert this vertical motion to forward thrust, giving it a characteristic saw tooth trajectory [17]. The opposite extreme of this low-energy glider is the ultrasonic thruster (UST) currently in development by Alfred Tan at MIT. By tuning an off-the-shelf ultrasonic transducer to a resonance frequency, a powerful jet can be obtained which can be used for propulsion. The revolutionary characteristic of the UST is that it has virtually no moving parts, a characteristic that lowers the maintenance cost and extends the lifespan of any machine [70].

Power

The capabilities and endurance of AUVs, like virtually all other vehicles without an integrated nuclear plant, are fundamentally limited by the amount of power available onboard. Both the total amount of energy available and the rate at which it can be discharged, i.e. power, must be considered. A comparison of several potential solutions is shown in Figure 4-2.

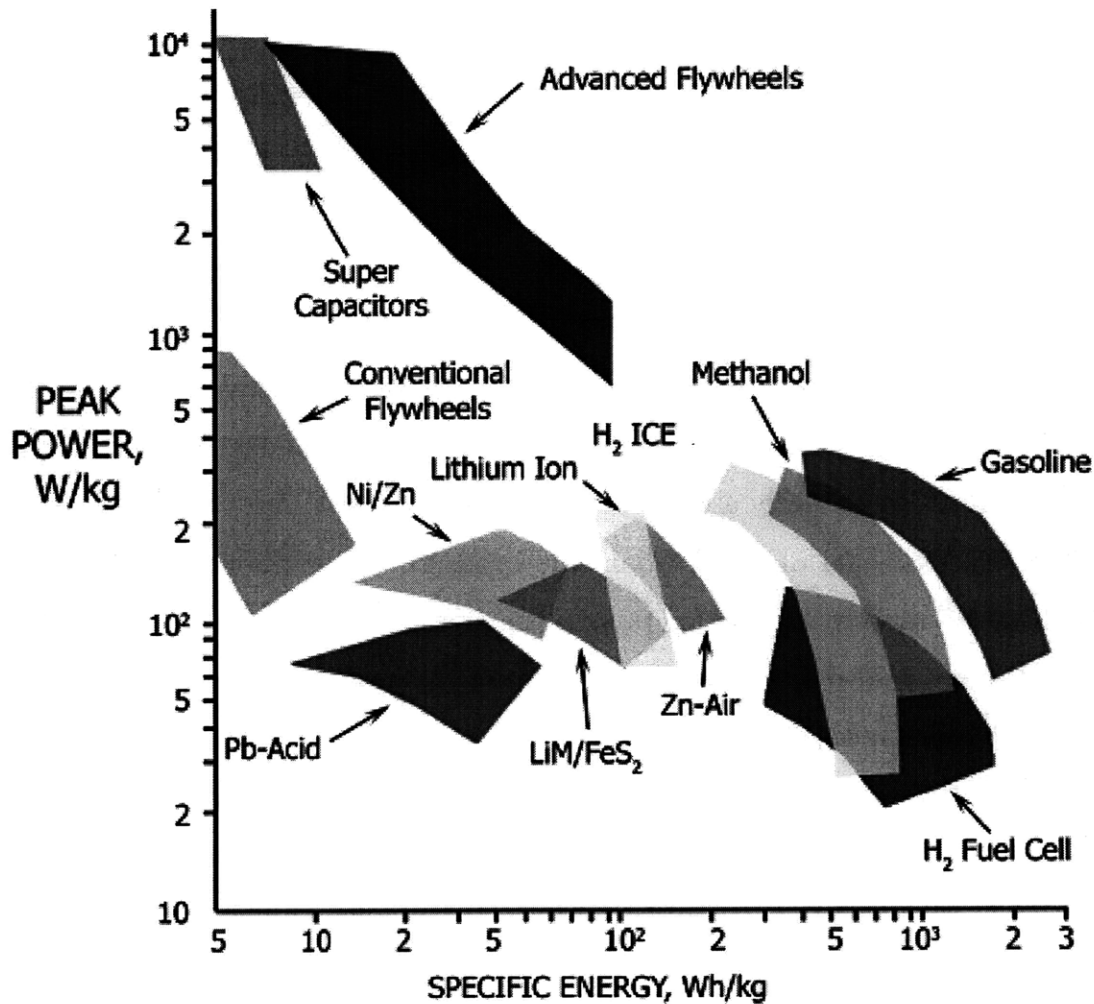


Figure 4-2: Comparison of power systems [29]

Traditionally, most AUVs have used either primary batteries, which can be used only once, or secondary batteries, which can be recharged [14]. The high cost of disposable primary batteries becomes economical only when the cost of the overall

AUV mission is particularly large, such as for scientific and defense applications [61]. Popular primary batteries include lithium-based cells, which are very powerful but equally expensive, and the common manganese alkaline cell. Silver-zinc cells have been used extensively as secondary batteries, but require careful monitoring and only last a few charging cycles. Lithium-solid polymer cells are pressure independent, a valuable characteristic in any AUV component [61]. Fuel cells can also be designed to be pressure-independent. For example, the Norwegian-developed HUGIN 3000 is equipped with an alkaline aluminum/hydrogen peroxide fuel cell. However, a drawback of this design is the necessity to reload the vehicle with potassium hydroxide and other potentially dangerous chemicals aboard a support ship [32].

Combustible fuels offer extremely high specific energy, but require an oxidizer to be carried onboard. Some batteries rely on oxygen dissolved in seawater, but the low concentration of oxygen invariably results in low voltage and hence low power density [14]. At the opposite extreme, super capacitors deliver relatively small amounts of energy very quickly. It is feasible that such technology could be used aboard an AUV to enable intervention missions, while an alternative source would power the thrusters. While flywheels are an extremely efficient method of storing energy, kinematic restrictions and the necessity for safety shielding limit their use aboard vehicles and they tend to only retain their “charge” for a few hours. In 1989 Stommel envisioned an AUV glider that derived its power directly from the sea. As the glider rises in warm waters, an internal bladder melts and expands against an accumulator that stores the energy for later use [17]. This technology would take full advantage of the unique opportunities offered by the maritime environment. However, while current gliders are extremely efficient, Stommel’s device has not yet been realized.

Launch and Recovery

The means by which an AUV is launched and recovered influence how delicate the payload may be, how easily actuators or other tools may be modified, how the vehicle can be refueled, and how recorded data is transferred from the vehicle to permanent databases. Most AUVs are currently both launched and deployed using a ship’s

crane, a very labor-intensive and sometimes dangerous process [10]. Adverse weather tends to affect recovery much more than launch, so there is greater opportunity to improve the process by concentrating on AUV recovery [31]. An alternative method of recovery is a stinger system in which the AUV drives into a cradle and catches a wire with a tail hook, similar to how airplanes land on an aircraft carrier [10]. ROVs are often lowered and hoisted in a garage, from which the ROV deploys underwater at operating depth [27]. It is feasible that such a system could be developed for AUVs.

Many engineers have envisioned a system of AUVs that are not regularly recovered by a ship, but rather dock to recharge their batteries and download mission data autonomously at stations. If such stations were floating on the surface, they could provide both solar power and GPS updates to the AUV's navigation systems [14]. Any such system would have to have an interface specially designed for the AUV in question. Garages, stingers, articulated arms, and nose cones could all be used to slide the AUV into place after it had successfully homed in on its base. However, careful attention must be paid to the AUVs power system when selecting a base location. Lead acid and certain other batteries give off dangerous, buoyant hydrogen gas when being recharged. Furthermore, the time spent by the AUV at its base station should be minimized to be economical in comparison to manually refueling the vehicle, but quick recharging can generate high temperatures [10]. Thus, the launch and recovery of the AUV is an important consideration when integrating together the vehicle's various subsystems.

Communication

AUVs rely on communication systems to navigate, to receive updates to their instructions, and to send data packets quickly. While the high conductivity of seawater attenuates all but the lowest-frequency radio waves, many AUVs are designed to partially breach the water's surface and deploy antennas to acquire a GPS signal. These vehicles normally communicate with the low-earth-orbit GPS satellites, which is more energy efficient than relying on higher-orbit satellites [17].

When submerged, AUVs rely on acoustic modems to communicate with one an-

other and with surface-based installations. Acoustic transmission typically occurs at 8.075 kHz and 27 kHz. The slow rate of acoustic transmissions creates a delay between any operator input and vehicle actions, increasing the need for autonomous processing aboard the AUV [14]. Solutions developed for terrestrial autonomous systems to mimic stigmergy in ant colonies by using the environment as a communication medium are inapplicable in the dynamic ocean. Many nations produce a wide array of acoustic communication systems for various applications. However, this profusion of options also means that acoustic channels can become cluttered in certain environments. For example, in Australia offshore platforms use acoustic beacons to warn submarines of their presence [27]. Acoustic signals are also susceptible to distortion and reflection off a shallow seabed or a group of ships, but remain the default communication option for undersea systems [14].

Command and Control

AUVs by definition require less operator control than other robotic systems. Increased artificial intelligence progresses a vehicle from simply being able to maintain position to executing preprogrammed tasks, such as opening a valve, to abstract mission-level assignments such as self-navigating to every valve in a platform and automatically checking their status and adjusting them as necessary. These tasks can be more complex than their terrestrial counterparts because moving manipulators and changing ocean conditions rapidly change the vehicle's basic dynamics during the mission [80]. One solution to such evolving conditions is the use of a computer supervisor that switches among multiple control schemes, depending on the situation [36]. In a layered control system, each level of task is handled separately [14]. When a layered control system is coupled with decentralized computing, where each subcomponent on board has the intelligence to handle its own particular tasks without much intervention from the central computer, the resultant design is very flexible towards future design changes and potentially produces much cost savings [3]. AUVs exhibit much potential for use in large groups, which would permit the application of swarm intelligence. The swarm would spontaneously coordinate activities to execute very

abstract missions with only the highest level of operator input.

Navigation

The navigation system aboard an AUV tracks the vehicle's position in space, typically at scales of several meters in resolution. While GPS provides excellent navigational data for other robots, many AUVs cannot surface frequently to receive a GPS signal. Ice or other structures may block their path, or a deep-sea AUV may not have the energy and time to expend resurfacing every time a new location update is desired. The simplest means of navigation is dead reckoning, i.e. estimating a vehicle's current location with knowledge of its starting position, velocity, and the time of travel. Since velocity measurements are prone to inaccuracy, significant improvement can be gained with the use of an inertial navigation system. These instruments use gyroscopes to record the acceleration of the AUV, and are available at many different levels of accuracy proportional to their cost. Such systems are often supplemented with Doppler velocity logs to record the vehicle's relative velocity [66]. A Doppler velocity log alone may be sufficient for navigation if it is able to lock onto the seafloor.

AUVs may also use acoustic networks to navigate. Long-baseline (LBL) networks consist of multiple transmitters, sometimes placed on the seabed up to ten kilometers apart. An AUV can determine its own location by recording how far it is from at least four of these transmitters, each of which has a known position [35]. An ultra-short baseline (USBL) system relies on an AUV's ability to determine the distance to a single transmitter mounted on the hull of a support ship. However, USBL methods require the use of a ship equipped with very accurate motion-sensing equipment [14] [22].

Theoretically, AUVs could navigate by locating geophysical features underwater. These features can either be naturally occurring or deliberately placed beforehand. However, practical obstacles to implementing such a system include difficulty of performing feature recognition from noisy sonar data. One solution to this challenge is vision-based navigation, but such a system would require laser illumination everywhere but the shallowest seas [15]. Simultaneous localization and mapping (SLAM)

techniques may be able to assist in AUV navigation if the vehicle revisits an area multiple times [66].

Positioning

Positioning is the capability of an AUV to locate and orient itself on a scale that is less than one meter in resolution. This feature is particularly important for intervention missions, on which an AUV interacts with relatively small systems such as valves or pipelines. While the accuracy of low-cost fluid tilt and pendulum tilt pitch and roll sensors degrades significantly in the presence of heave and sway, more expensive sensors are available that incorporate gyroscopes. High-cost gyroscopes are also available for use as angular velocity meters, but their cost and power demands have previously limited their use to military applications [39].

There is much overlap in the techniques and sensors used for navigation and positioning. Small-scale localization can be achieved by referencing known features using sonar or visual systems. A wide array of imaging sonar arrays are produced worldwide but an examination of their output shows that the resolution of these systems is too poor to distinguish small objects, such as features on oil rigs [27]. Machine vision algorithms can distinguish among even relatively similar features, and often can be retrofitted to systems originally deigned for human optical recognition [30]. However, such systems require computer processing and illumination that can be very power intensive, and would likely be very sensitive to biofouling.

Manipulators

Accurate positioning enables the use of manipulators on an AUV to interact with objects underwater. ROVs typically have manipulator arms preinstalled, and larger systems can lift hundreds of kilograms. While lower-cost manipulators have less dexterity and strength than more expensive alternatives, it has been speculated that the agility of the systems will eventually plateau and reach a cost constant but that the strength of the systems will always be proportional to cost [14]. However, it should be noted that each of the six degrees of freedom requires at least one actuator

and that sometimes more are used for redundancy or to increase the effectiveness of the arm. Generally, such systems will always be more complex and hence more expensive than their non-holonomic counterparts that are less agile.

While the most common manipulator is a simple gripper, many other tools can be affixed to a robotic arm. In the offshore industry, devices such as torque tools, gasket change-out tools, and pipe cutters have been used on ROVs [53]. The use of such manipulators on untethered systems has been rare because of the timing delay introduced by acoustic communication. However, as control algorithms improve it will become possible for AUVs to use increasing complex manipulators on their own. The feasibility of this scheme has already been demonstrated at the University of Hawaii with the SAUVIM (Semi Autonomous Underwater Vehicle for Intervention Missions), which can unplug, move, and reinstall a small part on an underwater frame on its own [45]. The ALIVE (Autonomous Light Intervention VEhicle) is capable of autonomously navigating to and docking at a work site before using a manipulator to perform light tasks. The experimental vehicle has explicitly been developed as a means to reduce the large costs associated with operating an ROV support ship [11].

Payload Sensors

An AUV may carry sensors unrelated to its actual motion or intervention missions. The litany of possible options currently available on the market include magnetometers, chemical sensors, conductivity sensors salinity sensors, thermometers, non-navigational pressure gauges, sediment profilers, radiation samplers, and plankton samplers [27]. These sensors are used to gather oceanographic data, to survey the seabed, or to conduct biological studies. Specialty forms of sonar, such as side-scan and bottom profiling, are also very frequently installed on underwater vehicles. The transceiver design is considered a mature technology in which only incremental improvements are possible, although research in onboard data compression and processing is ongoing [14]. Cameras are also available for a wide spectrum of costs, ranges, and lighting conditions [27]. However, the utility of any visual system is strongly correlated to the amount of power available for lighting.

4.2.3 Subsea Missions in the Offshore Industry

The offshore oil and gas industry has rapidly expanded since the 1960s, and provides up to a quarter of the natural gas in the United States and one-sixth of the nation's domestic oil [14]. The first offshore structures were simply steel towers, called jackets, that were fixed to the seafloor and supported the wellheads, pipes, and other fixtures beneath the petrochemical processing equipment and living quarters above the surface. As the offshore industry expanded into deeper waters tension-leg platforms (TLPs) and the floating production system (FPS) were developed. These configurations consist of the same topside facilities as a fixed structure, except that they are linked by flexible cables, rather than a rigid jacket, to the wellheads below [48]. This vital economic sector relies on underwater vehicles because the depths of modern offshore operations preclude the use of divers. While 83% of offshore activities occur at depths less than 300 m [16], the depth of drilling has increased with time so that the deepest structures now lie up to 2,400 m below the surface [58]. A wide variety of activities occur at these benthic depths, ranging in complexity from simple surveillance to the repair of damaged equipment.

Observation

The simplest task performed beneath an offshore rig is the visual inspection of equipment. Subsea installations can radiate out for up to 50 km beyond the oil rig seen on the surface. These pipes must be routinely inspected for damage [14]. Insurance regulators also sometimes mandate the visual inspection of equipment for cracks and other flaws according to a regular schedule. If such an investigation reveals a site in need of repair, an underwater vehicle can also be positioned to provide real-time video feed of the area while repairs are conducted with other tools on an ROV. Some observation tasks involve non-destructive testing, such as magnetic particle inspection to determine the structural integrity of a jacket node. Other observation tasks are more complex, and may include the spot removal of biological growth with a wire brush or other manipulative task before a measurement can be made [48].

Measurement

Measuring distances is a more complicated mission than simple observation because measurements must be made with reference to two or more points. This task requires a machine to briefly maintain a fixed position while establishing its position relative to both of the targets. One such measurement is of scour, the drifting of sand dunes along the seabed. The height of the sand is measured relative to a fixed point on the jacket, and can vary by as much as 4.5 m among the legs of the same jacket. Scour is important to monitor because it can expose the foundations of the jacket, posing a safety hazard [48].

Exploration

Exploration is the most sophisticated non-intervention task because such data is only useful if it is correlated to a specific geographical location. Deep-sea vehicles can map the sea floor using various instruments, such as multibeam echo sounders, sub-bottom profilers, side-scan sonar arrays, and magnetometers. It is significant that sonar-based measurements require much less data storage and processing than visual imagery. Such terrain data are critical when planning the installation of new subsea oil structures and reduce costly over-design of platforms. Interestingly, one of the first companies to invest largely in AUVs for exploration was the De Beers mining group, which was searching for diamonds off the coast of South Africa. De Beers contracted this work out to Maridan AS in Denmark [16].

Maintenance

After an offshore rig is built, adjustments must still be made in response to the evolving nature of the drilling operation. One such task is the opening and closing of valves on the "Christmas tree," the flow regulator which sits atop each wellhead on the sea floor [48]. Another routine assignment is the changing out of gaskets at various locations along the pipelines [53]. These tasks are not only simple and repetitive, but have the advantage that they require a deep-diving vehicle only to interface with

components predesigned for such interactions. Therefore, thoughtful planning prior to installation can make such maintenance much easier to perform onsite.

Construction

Engineers design offshore platforms so that as much of the structure as possible can be assembled on land before the components are sunk and connected underwater. Real-time monitoring is necessary to ensure that as parts are placed properly underwater. Careful observation and response is especially critical when laying pipeline in deep water, because the large loads frequently cause pipes to buckle and then rapidly collapse due to the high water pressure. Simple tasks, such as wire cutting, are also performed by underwater vehicles [16]. Technologically, it would be possible for robots to perform welding and other labor-intensive, high cost tasks undersea. However, regulators and insurance companies' strict standards require on-site human supervision and visual inspection of such joints and have so far precluded the use of robotic welders [59]. Although construction tasks are specialized one-time events in the lifespan of a rig, they can be anticipated and coordinated with the design of the component structures. Therefore, construction may be regarded as a hybrid task lying between intervention and maintenance.

Unique Tasks

The lifespan of offshore subsea equipment can extend up to 25 years, during which time any type of mechanical or electrical failure may occur. For example, a flaw in a piece of ExxonMobil equipment caused crystallization in a pipeline. In order to repair the damaged area, a custom ROV was designed and built to cut through the pipes surrounding the affected area on the seafloor and then bring the damaged segment to a surface vessel for further inspection [53]. The unpredictable nature of such repair tasks require very sophisticated decision making, and therefore are the most complicated assignments possible for a deep-diving vehicle.

Scientific Tasks

Offshore underwater missions are frequently encountering marine life by surprise. For example, in November 2007 a Shell ROV videotaped an “elbowed” *Magnapinna* squid, as shown in Figure 4-3. This was only one of a handful of sightings of the species in history [33].



Figure 4-3: *Magnapinna* squid photographed by a Shell ROV.

While such photographs frequently only reach biologists by circumlocutious routes after offshore employees save them for their novelty, formal relationships are being developed to match the technical capabilities of offshore deep-sea vehicles with the needs of scientists. In the United Kingdom, the SERPENT (Scientific and Environmental ROV Partnership using Existing iNdustry Technology) Project focuses on such unusual pairings [33]. While there is no technical need for offshore oil companies to participate in such projects, there may be large political and public-image incentives for a petrochemical company able to portray itself as being environmentally conscious and cognizant of and cooperative with cutting-edge scientific research.

Organization
Chevron
Chevron Energy Technology Company
International Submarine Engineering Ltd.
MIT Computer Science and Artificial Intelligence Laboratory
MIT Sea Grant
Monterrey Bay Aquarium Research Institute Engineering
Shell International Exploration and Production Inc.

Table 4.1: Organizations represented in the AUV design survey.

Therefore, since minimal additional capital investment may be necessary to accommodate biologists' needs, it would behoove an offshore company to consider scientific missions when designing future deep-sea vehicles.

4.2.4 Correlation between AUV Subsystems and Offshore Missions

Recall that the relationship between product subsystems and missions in the network-based design analysis are defined by the probability that a component will be pushed to peak capacity. This event infers that the mission could achieve more if the subsystem had more advanced capabilities. Collecting peak-capacity data for a novel system, such as an offshore AUV, is conjecture and limited by preconceptions. To help mitigate this bias, a survey has been conducted to collect the necessary information.

A questionnaire was written and distributed to a wide variety of specialists to collect data describing a potential AUV for the offshore industry. A copy of the survey has been included as Appendix A. An effort was made to contact members of the offshore oil industry, AUV companies, offshore ROV companies, and academia to create a robust data set. A list of participating organizations is given in Table 4.1.

General Survey Responses

One major limitation of the survey was the presumption that a single type of AUV would conduct all of the missions considered. This simplification was made to avoid assumptions with regards to the most efficient way of grouping the tasks assigned to different types of AUVs. However, many of the survey respondents felt that it was more efficient for a corporation to purchase multiple specially-built AUVs for each mission type. The survey respondents also felt that it would be more practical to design unique AUVs for separate geographical conditions rather than select a single omnibus vehicle. For example, very different properties would be demanded by operations underneath an ice shelf, in extraordinarily deep water, and in coastal areas.

Several other general trends emerged from the survey results. The respondents felt that the capabilities of any AUV would be severely limited by its power supply, and that this component was inherently much more influential than any other subsystem. Several respondents also noted their belief that AUVs would not have the technical capabilities to conduct construction or unique intervention missions for at least another ten years.

The information gathered in the survey and associated conversations is presented in Table 4.2. This data does not represent the viewpoint of any one participant, but rather is a compilation of the opinions expressed by all of the respondents. Note that whether a subsystem reaches peak capacity in different missions are independent events. Therefore, the probabilities of a given subsystem reaching peak capacity in the various missions need not sum to 100%. The peaking of different subsystems during the same mission are not mutually exclusive events, so the sum of the probabilities of each subsystem peaking during the same mission is also unconstrained.

Explanation of Survey Results

The respondents' predicted division of AUV service time was very interesting. Their enthusiasm towards an AUV performing observation tasks may have seemed surpris-

	Observation	Measurement	Exploration	Maintenance	Construction	Unique	Scientific
Significance	28%	17%	17%	14%	8%	7%	9%
Materials	0	0	25%	30%	35%	30%	16%
Propulsion	20%	15%	25%	30%	32%	25%	25%
Power	50%	60%	60%	75%	80%	65%	60%
Launch and Recovery	18%	20%	25%	25%	30%	25%	22%
Communication	23%	15%	38%	42%	50%	38%	24%
Command and Control	14%	25%	21%	50%	50%	46%	20%
Navigation	30%	21%	53%	35%	26%	30%	35%
Positioning	28%	32%	17%	36%	44%	33%	18%
Manipulators	0	0	0	55%	60%	55%	15%
Payload Sensors	22%	28%	28%	15%	15%	15%	70%

Table 4.2: The correlation between AUV subsystems and offshore oil missions.

ing, considering that in such a role the vehicle would be little more than a floating camera and that significant operator time would still be required to perform the main task at hand which the AUV was observing. However, such tasks can be complicated by factors such as clutter, operating depth, noise, and currents. While AUVs have frequently been cited as ideal exploration vehicles, one survey participant emphasized that the task could be fulfilled sufficiently by a surface craft and towed array. Multiple respondents suggested that intervention tasks such as construction and unique missions could perhaps be best performed by a hybrid ROV/AUV.

The materials comprising an AUV were the primary determinant of what regions the vehicle could access safely. The strong correlation between a vehicle's materials and exploration capabilities was therefore straightforward. Somewhat more surprising was the opinion shared by many respondents that intervention missions, particularly construction, would be limited by the quality of materials. However, as one participant suggested, vehicles engaged in such heavy-duty tasks would need to be larger and more robust than their counterparts that engaged strictly in passive observation.

Many of the respondents, particularly those from the offshore industry, felt the

maneuvering capabilities of propulsion subsystems would be taxed by construction and maintenance work. Some other survey participants appeared to focus more on the possibility that propulsion systems would limit the range of exploration missions. In practice, however, it is possible that a vehicle's range would instead be limited more by its power capabilities. In the case of exploration, sometimes the range of the vehicle could be extended significantly by implementing innovative autonomous recharging sites. The peak energy output rates in particular are especially critical when performing intervention tasks.

There was a consensus among the survey participants that the launch and recovery of a vehicle rarely limited mission capabilities. The exception to this general rule was when subsea operations were conducted in especially remote locations. The communication abilities of an AUV were inherently limited by the ocean environment, and vehicles would be forced to surface to achieve large bandwidths. The intensity of communication was believed to increase with the complexity of the AUV mission. Command and control demands were also believed to scale with the complexity of the mission.

The survey participants perceived navigation and positioning as cooperative subsystems that would reach peak performance in tandem. The only exception to this trend is that navigation is more critical when the vehicle is very far from established subsea structures, such as in exploration and construction missions. Manipulators were also deemed to be most important in a small subset of missions, those in which the AUV must interact with its environment. However, one respondent emphasized that manipulators were very versatile because a single claw could grip a multitude of tools. This adaptability limited the probability of manipulators limiting offshore missions. In contrast to manipulators, payload sensors were most important in missions that focused more on observation than intervention. For example, the sensors that happened to be aboard an AUV would limit the incidental scientific missions that the vehicle could undertake. However, exploration, observation, and measurement missions were deemed less likely to be limited by the sensors aboard. This conclusion may have been because AUVs explicitly defined for such activities would have a high

Subsystem	Flux Value
Power	2.686
Manipulators	0.354
Navigation	0.138
Communication	-0.083
Command and Control	-0.276
Positioning	-0.282
Materials	-0.476
Payload Sensors	-0.486
Propulsion	-0.743
Launch and Recovery	-0.833

Table 4.3: Flux ranking of AUV subsystems.

degree of built-in redundancy and would not demand peak behavior from the sensor systems.

4.2.5 Significance of Subsystems in AUV Design

Applying Equation 4.3 to the data in Table 4.2 yields the flux value of each subsystem. These scores are shown in Table 4.3.

Recall from Section 4.1.1 that the flux value of a subsystem corresponds to the probability that it reaches peak capacity and therefore limits mission capabilities. It is logical that the power subsystem would have the greatest flux value; this was the component most frequently cited by the panel of experts as a limiting factor in AUV design. The subcomponents with the positive flux scores, i.e. power, manipulators, and navigation, are more likely to peak than the other subsystems involve in the same missions. Another interesting result is that the lower-ranked subsystems are those components that seem to be very well developed, such as materials or launch and recovery. A common trait of these components is that they are not being actively developed in the marketplace. This lack of attention may reflect a desire to concentrate on projects that yield the greatest opportunity to increase profit by expanding mission capabilities.

Mission	Flux Value
Observation	1.490
Exploration	0.714
Maintenance	0.683
Measurement	0.042
Construction	-0.690
Scientific	-1.001
Unique	-1.237

Table 4.4: Flux ranking of offshore missions.

4.2.6 Significance of Missions in AUV Design

Applying Equation 4.6 to the data in Table 4.2 yields the flux value of each mission, which are shown in Table 4.4. The flux value of a mission corresponds to the likelihood that it will drive subcomponents to their peak capacity as compared to other missions that share those subcomponents. A large positive flux score may indicate that mission requirements are too stringent for the current system configuration. As stated in Section 4.1.2, a greater flux score for a mission is associated with a larger share of maintenance costs due to increased wear and tear. It is interesting, therefore, that according to Table 4.4 greater maintenance costs are associated with science missions, which are ancillary activities to the operation of the offshore platform, than are generated by unique missions, which normally occur during emergency repairs. The non-intervention missions in Table 4.4 would rely on fewer technological advances than the intervention missions to be conducted autonomously. The non-intervention missions could therefore be expected to have a lower capital cost than the intervention missions, such as construction. However, the study results indicate that this cost savings will be counterbalanced by a greater share of the maintenance cost being incurred by the non-intervention missions, which have positive flux scores. It should be noted that the relative ranking of missions by flux value is not identical to the ranking of the missions according to their significance. This phenomenon occurs because missions that occupy less of an AUV's total operation time can still place greater strain on AUV subsystems than more frequent tasks. This insight is one

reason that the flux metric is a valuable tool to a corporation analyzing the true costs associated with AUV operations.

4.3 Advantage of Flux Metric in Design

The use and interpretation of the flux metric in a design setting has been demonstrated, but it may not be clear what advantage this technique offers to an engineer. Traditional quantitative design techniques, while not explicitly based on network theory, already share similar properties. Furthermore, even if the application of network methods to engineering design is deemed beneficial it may not be clear which network metric is the most appropriate analysis tool. However, it can be shown that the flux metric offers benefits distinct from those of alternative approaches to structured design.

4.3.1 Comparison to House of Quality

Although the conceptualization of the design space as a network is apparently unique, the calculations that result bear a strong resemblance in spirit to the construction of a house of quality and other tools of the quality function deployment method (QFD) described in Ullman and other design textbooks [73]. However, the insight gained by the application of the flux metric is more robust than QFD, and thus advances beyond those traditional techniques in certain situations.

The house of quality (HOQ) is the main tool of QFD. The principle of the house of quality is to correlate the customers' desires with engineering specifications. For example, in the example from Section 4.1, the HOQ score of subsystem A would be

$$HOQ(A) = P_{\alpha}A_{\alpha} + P_{\beta}A_{\beta} \quad (4.7)$$

A comparison of Equation 4.3 and Equation 4.7 reveals that the HOQ score is equivalent to the weighted outdegree of the subsystem, but does not account for the number of other subsystems utilized in the same mission. In the case of the offshore

Subsystem	House of Quality Score
Power	0.613
Navigation	0.332
Communication	0.302
Positioning	0.287
Command and Control	0.278
Payload Sensors	0.263
Propulsion	0.232
Launch and Recovery	0.223
Manipulators	0.177
Materials	0.148

Table 4.5: House of Quality ranking of AUV subsystems.

AUV case study, the house of quality calculation would be conducted by multiplying the correlation between a subsystem and each mission in Table 4.2 with that mission’s significance and then summing the results for each subsystem. The resulting values are shown in Table 4.5.

The drawback of the house of quality method in comparison to flux is that the house of quality considers each subsystem independently of the existence of the other subsystems. Therefore, the house of quality assigns a high score to those subsystems that relate to a large number of missions even if the relationship between the component and the task is relatively weak in comparison to the other subsystems associated with that objective. For example, the house of quality methods ranks positioning as an especially important subsystem even though the probability of that subcomponent peaking during a mission is never greater than 45%. Note that if the matrix describing the relationships between the subsystems and missions is fully connected, i.e. every subsystem is utilized by every mission, then the ranking according to the house of quality will be identical to the flux ranking. This special case occurs because when a fully-connected bipartite graph is decomposed to its unipartite form every node has the same indegree. In a fully-connected graph the outdegrees of the subsystems are equal to their house of quality scores multiplied by the total number of subsystems.

4.3.2 Comparison to Other Network Metrics

The principles developed in this chapter establish the application of network methods to engineering design. This infers that theoretically any network metric could be effectively applied to a design problem. This is technically true, but an engineering network bears all of the complexity of more traditional networks and the precept remains that there is no omnibus tool that may be applied without first carefully examining its conceptual assumptions and implications. Among the centrality measures examined in this project, flux remains the most appropriate measure to be applied to design networks.

It is clear that an unweighted design network would contain much less information than its weighted counterpart. As Table 4.2 demonstrates, such a network may be fully-connected or nearly so and thus would be meaningless without appropriate edge weightings. If a model depicts components that support or influence one another, as does the technique presented in this chapter, then directed edges should also be used. These two conditions infer that degree and betweenness would be inappropriate for use in this type of design network. However, it should be noted that if only the correlation between two subsystems is known then an undirected graph may be more appropriate and a greater variety of network metrics would be applicable.

The only network metric appropriate for weighted, directed networks discussed in this project other than flux is eigenvector centrality. However, this tool is built on the premise that a node gains importance by being linked to other important nodes. In the case of a design network it is not clear if this premise would apply. The only relevant relationship between subsystems is the relative likelihood that each subcomponent will be driven to peak capacity during the same mission. There is no intuitive significance to a subsystem reaching peak capacity in a mission that it shares with a subcomponent that frequently peaks during another mission. Therefore, the flux metric is the most appropriate network analysis tool for design metrics.

Conclusion

The interpretation of an engineering design challenge as a complex network permits the application of network analysis methods to this class of problems. These techniques allow corporations to better understand the costs associated with system tasks and how various subsystems influence the capabilities of the overall product. The application of these techniques to the design of AUVs for the offshore oil industry has been demonstrated. Framing this design challenge as a network problem permits the use of nodal analysis techniques to identify the AUV subcomponents whose development should be prioritized to increase the quality of the total product. This analysis has identified an AUV's power, manipulators, and navigation subsystems as the components that are most likely to reach peak capacity and limit the success of a mission. The advantage of the flux metric in comparison to other potential network techniques for engineering design also has been explored.

Chapter 5

Conclusion

This project has introduced the flux metric, a measure of influence in a network that successfully differentiates among nodes that other centrality metrics regard as identical. The flux metric offers insight into the relationships among the subsystems and objectives of an engineered product and other complex systems.

5.1 Summary

Chapter 2 introduces the flux metric as a method to rank the influence of nodes in a weighted, directed network. Flux resolves the limitations of earlier centrality metrics, such as eigenvector centrality. The Secretary Paradox emphasizes the misleading results that such methods may yield, especially if nodes have equal weighted indegree values.

The Conserved Flow models demonstrates the applicability of flux to graphs of varying architecture, such as trigonal and rectangular, and of different sizes. In every situation, the flux of a perturbation site consistently indicates the rate at which the system will reestablish an equilibrium. The concept of flux neighborhoods further refine the interpretation and proper application of this novel metric.

In Chapter 3 the flux metric is used to analyze more networks from engineering and the social sciences. The identical units of these networks gain their varying properties from their orientation with respect to their neighbors. The Synchronization model

represents the spread of a shock through a network of opinions. The greater the flux value of the perturbation site, the greater the impact of the shock on the state values of the nodes throughout the network.

The Infection model simulates the life cycle of an epidemic on a network. This probabilistic system is an extension of the differential equations that traditionally are used to model an outbreak. The influence of a node in an Infection network is diminished by probabilistic effects, which renders the flux metric less insightful than it is in deterministic cases.

Resource Sharing simulates the distribution of a conserved good. The flux value of a node influences both the speed at which material will be dispersed from that site and the steady-state amount present there. The behavior of the network can be influenced by adjusting the greediness of each node during their exchanges.

Chapter 4 presents the concept of an engineering design challenge as a network problem. This premise is demonstrated by a case study analyzing the potential design of autonomous underwater vehicles (AUVs) for the offshore oil and gas industry. A review of developing technologies establishes the relevant AUV subsystems and offshore missions, and an industry-wide survey determines the correlation among these components. The flux metric offers insight into the design of a system that is not provided by other quantitative design techniques.

5.2 Future Work

Network studies are a burgeoning area of science and engineering that offer many opportunities for continued research and the potential discovery of valuable applications. Although the flux metric is capable of ranking the influence of nodes that other centrality measures treat as identical entities, flux does not appear to perform well in probabilistic settings and very likely has other limitations not exposed in this project. Therefore, the quest continues for a single technique that systematically combines the advantages of flux and the other centrality measures into a single algorithm.

Furthermore, the characteristics of the flux metric derived from computer sim-

ulations must be verified in empirical studies. However, without an independent definition of influence flux and all other centrality measures rely on intuition to validate their output. One potential solution is to use a network method to rank the influence of members of the United States House of Representatives using a bipartite graph that describes their membership in committees and caucuses and to then compare these results to influence measured by another technique, such as fundraising dollars or successful votes.

Recall from Section 2.1 that flux is defined as the difference between $\phi(\textit{outdegree})$ and $\phi(\textit{indegree})$. In all of the examples presented in this project the function ϕ is simply the identity function. However, ϕ could be defined by other functions. For example, if ϕ were a logarithm then the ratio of the outdegree to the indegree could be analyzed. The definition of ϕ should be carefully selected for a given application to match qualitative models of the system or other available calibration data.

The concept of an engineering design challenge as a network problem presented in this project could be developed further. For example, there are many valid interpretations of edge weightings other than the probability of peak behavior. Some design networks may be derived directly from a unipartite graph and lack directed edges. Network techniques could be applied to other business and management decisions, such as identifying the most critical steps in a flowchart.

The flux metric potentially can contribute to scientists' understanding of the many complex networks encountered in engineering and the social sciences. Hopefully flux itself will become an influential component in the network of techniques that we use to study and improve our world.

Appendix A

AUV Design Survey

The following survey was distributed to members of the offshore and AUV communities by email in June and July, 2009. A list of respondents has been included as Table 4.1.

Instructions

The objective of this survey is to identify which AUV subsystems would be pushed to their design capacity during various types of missions in the offshore oil industry. Examples of subsystems reaching this limit would be when Materials are at their rated depth or temperature limit, or when Propulsion systems are exerting maximum thrust.

The subsystems are:

Materials: The structural components of the AUV, such as the hull and supports.

Propulsion: The thrusters that drive and steer the AUV.

Power: The energy available onboard to power other subsystems, such as thrusters and instruments.

Launch & Recovery: The means by which the AUV is deployed, retrieved, or docked.

Communication: Data transfer between the AUV and any other database, whether by radio, satellite, or physical media.

Command & Control: The interface between operators and individual AUVs or teams of AUVs.

Navigation: How the AUV determines its geographic location.

Positioning: How the AUV determines its position on a local scale.

Manipulators: Arms, grippers, and tools with which the AUV affects its environment.

Payload Sensors: Instruments carried aboard an AUV that do not relate to the motion of the vehicle.

The offshore oil missions are:

Observation: The visual inspection of underwater equipment.

Measurement: Determining the distance between two objects or collecting other data.

Exploration: Searching unknown territory for new construction sites or resources.

Maintenance: Adjusting valves and other simple, repetitive tasks performed on preselected mechanical interfaces underwater.

Construction: The underwater assembly of predesigned structural components.

Unique: Unpredictable tasks, such as repairs, that require the vehicle to execute novel operations.

Scientific: Biological and geological research undertaken by offshore oil companies to share resources and foster their relationship with the scientific community.

Each subsystem is listed below, followed by seven sets of parentheses labeled with mission types. Please type directly in the appropriate parentheses the probability in percent that the subsystem reaches its capacity during that mission. Please only consider whether or not a subsystem reaches its limit during the mission, not how long capacity-level operation is endured. For example, you may feel that Propulsion subsystems would be pushed to capacity during only 25% of

Observation missions. It may be helpful to consider which subsystems capabilities would limit the performance of an AUV on a certain mission. If you do not believe that a subsystem would be used at all during the mission, then please type 0 in the parentheses.

Materials: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Propulsion: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Power: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Launch & Recovery: Observation() Measurement() Exploration()
Maintenance() Construction() Unique() Scientific()

Communication: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Command & Control: Observation() Measurement() Exploration()
Maintenance() Construction() Unique() Scientific()

Navigation: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Positioning: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Manipulators: Observation() Measurement() Exploration() Maintenance()
Construction() Unique() Scientific()

Payload Sensors: Observation() Measurement() Exploration()
Maintenance() Construction() Unique() Scientific()

Bibliography

- [1] Ian F. Akyildiz, Dario Pompili, and Tommaso Melodia. Underwater acoustic sensor networks: Research challenges. *Ad Hoc Networks*, 3:257–279, 2005.
- [2] David C. Bell, John S. Atkinson, and Jerry W. Carlson. Centrality measures for disease transmission networks. *Social Networks*, 21:1–21, 1999.
- [3] Anders Bjerrum. Auvs for surveys in complex environments. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 203–216. Taylor & Francis, New York, 2003.
- [4] Phillip Bonacich. Power and centrality: A family of measures. *The American Journal of Sociology*, 92(5):1170–82, March 1987.
- [5] Phillip Bonacich. Some unique properties of eigenvector centrality. *Social Networks*, 29:555–564, 2007.
- [6] Phillip Bonacich and Paulette Loyd. Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23:191–201, 2001.
- [7] Stephen P. Borgatti. Centrality and network flow. *Social Networks*, 27:55–71, 2005.
- [8] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30:107–117, 1998.
- [9] Vincent Buskens. The social structure of trust. *Social Networks*, 20:265–289, 1998.
- [10] Robin Galleti Di Cadilhac and Attilio Brighenti. Docking systems. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 93–108. Taylor & Francis, New York, 2003.
- [11] European Commission. Alive: New autonomous vehicle tackles expanding undersea applications. http://ec.europa.eu/research/transport/news/article_747_en.html, 2004.
- [12] Simon Corfield and Christopher Hillenbrand. Defence applications for unmanned underwater vehicles. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 161–178. Taylor & Francis, New York, 2003.

- [13] Bluefin Robotics Corporation. Homepage. [http://www.bluefinrobotics.com /index.htm](http://www.bluefinrobotics.com/index.htm), April 2009.
- [14] National Research Council. *Undersea Vehicles and National Needs*. National Academy Press, Washington, 1996.
- [15] F. Dalglish, S. Tetlow, and R.L. Allwood. Seabed-relative navigation by hybrid structured lighting. In Geoff Roberts and Robert Sutton, editors, *Advances in Unmanned Marine Vehicles*, pages 277–292. The Institution of Electrical Engineers, Stevenage, UK, 2006.
- [16] Edwin F.S. Danson. Auv tasks in the offshore industry. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 127–138. Taylor & Francis, New York, 2003.
- [17] Russ E. Davis, Charles C. Eriksen, and Clayton P. Jones. Autonomous buoyancy-driven underwater vehicles. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 37–58. Taylor & Francis, New York, 2003.
- [18] Morris H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345), March 1974.
- [19] Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic algorithms for replicated database maintenance. *Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing*, pages 1–12, 1987.
- [20] Reinhard Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer Berlin Heidelberg, Reading, Massachusetts, third edition, 2006.
- [21] Luca Donetti, Pablo I. Hurtado, and Miguel A. Muñoz. Entangled networks, synchronization, and optimal network topology. *Physical Review Letters*, 95, October 2005.
- [22] International Submarine Engineering. Auv design info page. <http://www.ise.bc.ca/WADENavandpos.html>, 2000.
- [23] J. Alexander Fax and Richard M. Murray. Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49(9), September 2004.
- [24] James H. Fowler. Connecting the congress: A study of cosponsorship networks. *Political Analysis*, 14:456–487, 2006.
- [25] Linton C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1997.

- [26] Samuel R. Friedman. Promising social network research results and suggestions for a research agenda. In Richard H. Needle, Susan L. Coyle, Sander G. Genser, and Robert T. Trotter II, editors, *Social Networks, Drug Abuse, and HIV*, number 151 in National Institute on Drug Abuse Research Monograph Series. National Institutes of Health, Rockville, MD, 1995.
- [27] Clifford Funnell, editor. *Jane's Underwater Technology*. Jane's Information Group, Coulsdon, UK, sixth edition, 2003.
- [28] Azra C. Ghani, Jonathan Swinton, and Geoff P. Garnett. The role of sexual partnership networks in the epidemiology of gonorrhoea. *Sexually Transmitted Diseases*, 24, 1997.
- [29] Ahmed F. Ghoniem. 2.62 fundamentals of advanced energy conversion lecture 23-2: Solar energy and solar thermal electric systems. Massachusetts Institute of Technology, Cambridge, MA, May 2009.
- [30] Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins. *Digital Image Processing Using MATLAB*. Pearson Prentice Hall, Upper Saddle River, NJ, 2004.
- [31] Gwyn Griffiths, Nicholas S. Millard, and Roland Rogers. Logistics, risks and procedures concerning autonomous underwater vehicles. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 279–294. Taylor & Francis, New York, 2003.
- [32] Øistein Hasvold, Nils J. Størkersen, Sissel Forseth, and Torleif Lian. Power sources for autonomous underwater vehicles. *Journal of Power Sources*, 162, 2006.
- [33] Kelly Hearn. National geographic news: Alien-like squid with ‘elbows’ filmed at drilling site. <http://news.nationalgeographic.com/news/2008/11/081124-giant-squid-magnapinna.html>, November 2008.
- [34] Herbert W. Hethcote. The mathematics of infectious diseases. *Society for Industrial and Applied Mathematics Review*, 42(4):599–653, 2000.
- [35] Hydro International. Hydro surveys: Long baseline systems. *Hydro International*, January 2008.
- [36] G. Ippoliti, L. Jetto, and S. Longhi. Switching-based supervisory control of underwater vehicles. In Geoff Roberts and Robert Sutton, editors, *Advances in Unmanned Marine Vehicles*, pages 105–126. The Institution of Electrical Engineers, Stevenage, UK, 2006.
- [37] Michael V. Jakuba. Modeling and control of an autonomous underwater vehicle with combined foil/thruster actuators. Master's thesis, Massachusetts Institute of Technology, 2003.

- [38] W.O. Kermack and A.G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, 115, 1927.
- [39] James C. Kinsey, Ryan M. Eustice, and Louis L. Whitcomb. A survey of underwater vehicle navigation: Recent advances and new challenges. In *Proceedings of the IFAC Conference of Manoeuvring and Control of Marine Craft*, September 2006.
- [40] MIT Sea Grant AUV Lab. Wiki. http://seagrant.mit.edu/auvwiki/index.php/Main_Page, March 2009.
- [41] Joel H. Levine. The sphere of influence. *American Sociological Review*, 37:14–27, 1972.
- [42] JS Liu, W-M Lu, C Yang, and M Chuang. A network-based approach for increasing discrimination in data envelopment analysis. *Journal of the Operational Research Society*, 1(9), 2009.
- [43] Jan Lorenz. Consensus strikes back in the hegselmann-krause model of continuous opinion dynamics under bounded confidence. *Journal of Artificial Societies and Social Simulation*, 9(1), 2006.
- [44] Zeev Maoz, Lesley G. Terris, Ranan D. Kuperman, and Ilan Talmud. International relations: A network approach. In Alex Mintz and Bruce Russett, editors, *New Directions for International Relations*. Lexington, Lanham, MD, 2005.
- [45] Giacomo Marani, Song K. Choi, and Junku Yuh. Underwater autonomous manipulation for intervention missions auvs. *Ocean Engineering*, 36:15–23, 2009.
- [46] K. Masabuchi, K. Itoga, J. Symonds, E.C. Hobaica, F. Campisi, S. Penley, J. Plummer, and V.T. Boatwright. Materials. In E. Eugene Allmendinger, editor, *Submersible Vehicle Systems Design*, pages 109–190. The Society of Naval Architects and Marine Engineers, Jersey City, NJ, 1990.
- [47] Naoki Masuda and Norio Konno. Transmission of severe acute respiratory syndrome in dynamical small-world networks. *Physical Review E*, 69, 2004.
- [48] Angus Mather. *Offshore Engineering: An Introduction*. Witherby & Company Limited, London, second edition, 2000.
- [49] Luc Moreau. Leaderless coordination via bidirectional and unidirectional time-dependent coordination. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003.
- [50] J.D. Murray. *Mathematical Biology*, volume 1. Springer, New York, third edition, 2002.
- [51] Mark Newman. Analysis of weighted networks. *Physical Review*, 70(5), 2004.

- [52] Office of Naval Research. Science & technology focus: Ocean water pressure. <http://www.onr.navy.mil/Focus/ocean/water/pressure1.htm>, April 2009.
- [53] Meghan Overstake. Subsea robotics in ExxonMobil's offshore operations. Massachusetts Institute of Technology Department of Mechanical Engineering Seminar. Massachusetts Institute of Technology, Cambridge, MA, April 2009.
- [54] Juan G. Restrepo, Edward Ott, and Brian R. Hunt. Characterizing the dynamical importance of network nodes and links. *Physical Review Letters*, 97, 2006.
- [55] Richard Rothenberg and Jerry Narramore. Commentary: The relevance of social network concepts to sexually transmitted disease control. *Sexually Transmitted Diseases*, 23, 1996.
- [56] Richard B. Rothenberg, John J. Potterat, Donald E. Woodhouse, William W. Darrow, Stephen Q. Muth, and Alden S. Klovdahl. Choosing a centrality measure: Epidemiological correlates in the Colorado Springs study of social networks. *Social Networks*, 17:273–297, 1995.
- [57] Britta Ruhnau. Eigenvector-centrality – a node-centrality? *Social Networks*, 22:357–365, 2000.
- [58] Rich Sears. Oil, natural gas and energy systems. Massachusetts Institute of Technology 2.62 Lecture. Shell Exploration & Production Ltd., The Hague, March 2009.
- [59] Randive Inc. Commercial Diving Services. Personal interview. Connecticut Maritime Association, Inc. Shipping 2009, March 2009.
- [60] Suleiman M. Abu Sharkh. Propulsion systems for auvs. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 109–126. Taylor & Francis, New York, 2003.
- [61] Suleiman M. Abu Sharkh, Gwyn Griffiths, and Andrew T. Webb. Power systems for unmanned underwater vehicles. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 19–36. Taylor & Francis, New York, 2003.
- [62] Hanumant Singh, James G. Bellingham, Franz Hover, Steven Lerner, Bradley A. Moran, Keith von der Heydt, and Dana Yoerger. Docking for an autonomous ocean sampling network. *IEEE Journal of Oceanic Engineering*, 26(4):498–514, October 2001.
- [63] F. Slanina, K. Sznajd-Weron, and P. Przybyla. Some new results on one-dimensional outflow dynamics. *EPL*, 82, 2008.

- [64] Peter Stevenson and Derek Graham. Advanced materials and their influence on the structural design of auvs. In Gwyn Griffiths, editor, *Technology and Applications of Underwater Autonomous Vehicles*, pages 77–92. Taylor & Francis, New York, 2003.
- [65] E. Stump, A. Jadbabaie, and V. Kumar. Connectivity management in mobile robotic teams. *IEEE International Conference on Robotics and Automation*, May 2008.
- [66] Luke Stutters, Honghai Liu, Carl Tiltman, and David J. Brown. Navigation technologies for autonomous underwater vehicles. *IEEE Transactions on Systems, Man, and Cybernetics*, 38(4), July 2008.
- [67] Jun Sun, Stephen Boyd, Lin Xiao, and Persi Diaconis. The fastest mixing markov process on a graph and a connection to a maximum variance unfolding problem. *Society for Industrial and Applied Mathematics Review*, 48(4):681–699, 2006.
- [68] K. Sznajd-Weron. Controlling simple dynamics by a disagreement function. *Physical Review E*, 66, 2002.
- [69] Katarzyna Sznajd-Weron and Jozef Sznajd. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(6):1157–1165, 2000.
- [70] Alfred Tan and Franz Hover. Ultrasonic thruster (ust). Poster. Cambridge, MA: Massachusetts Institute of Technology, 2008.
- [71] Michael Taylor. Influence structures. *Sociometry*, 32(4):490–502, December 1969.
- [72] Chris Allen Thomas. Centrality in networks. <http://www.mysocialnetwork.net/blog08/555/r3008/2008/02/centrality-in-networks.html>, February 2008.
- [73] David G. Ullman. *The Mechanical Design Process*. McGraw-Hill, New York, third edition, 2003.
- [74] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge University Press, Cambridge, UK, 1994.
- [75] Duncan Watts and Steven Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:409–10, June 1998.
- [76] Duncan J. Watts. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press, Princeton, NJ, 1999.
- [77] Douglas R. White and Stephen P. Borgatti. Betweenness centrality measures for directed graphs. *Social Networks*, 16:335–346, 1994.
- [78] Sarita Yardi and Amy Bruckman. Modeling the flow of information in a social network. *ACM SIGCOMM 2008*, 2008.

- [79] Jennifer Yick, Biswanath Mukherjee, and Dipak Ghosal. Wireless sensor network survey. *Computer Networks*, 52:2292–2330, 2008.
- [80] J. Yuh. Design and control of autonomous underwater robots: A survey. *Autonomous Robots*, 8:7–24, 2000.