MONASH University

Understanding the Deception Pandemic

Dr Carlo Kopp, Fellow LSS, A/Fellow AIAA, SMIEEE, PEng
Faculty of Information Technology, Monash University, Clayton, 3800

GROUP OF EIGHT AUSTRALIA

About the Author

- Part time computer science academic at Monash University Faculty of IT since 2000
- ~680 publications in multiple categories since 1979
- http://users.monash.edu/~ckopp/hypsys.html
- Best known as military analyst and strategist since 1979, Co-founder of *Air Power Australia* military strategy think-tank @ http://www.ausairpower.net/
- In 1999 co-discovered the *"Borden-Kopp"* information theoretic models of deception

MONASH University

The Deception Pandemic

Deception + Pandemic - Wikipedia

- **“Deception** is the act of propagating a belief that is not true, or is not the whole truth (as in half-truths or omission). Deception can involve dissimulation, propaganda, and sleight of hand, as well as distraction, camouflage, or concealment. There is also self-deception, as in bad faith. It can also be called, with varying subjective implications, beguilement, deceit, bluff, mystification, ruse, or subterfuge.”
- “A **pandemic** (from Greek πᾶν *pan* "all" and δῆμος *demos* "people") is an epidemic of infectious disease that has spread across a large region; for instance multiple continents, or even worldwide.”

Observation

- 500 years ago Gutenberg's printing press played a key role in the *Reformation*, employed to print large numbers of *"Reformation Pamphlets"*, i.e. *propaganda* to promote Church reform[1].
- This is an important case study – changing means of mass distribution of information were first exploited for political and social purposes, reflecting the cultural foci of that period.
- The use of digital media for the mass distribution of political and commercial propaganda, *"fake news"*, *"clickbait"* and every other form of dis/mis-information follows much the same pattern today, but with contemporary cultural foci.

1 Bagchi D. Printing*, Propaganda, and Public Opinion in the Age of Martin Luther*; 2016. Oxford Research Encyclopedia of Religion. http://religion.oxfordre.com/view/10.1093/acrefore/9780199340378.001.0001/acrefore-9780199340378-e-269

Mapping the Deception Pandemic

- Social and Mass Media – the *"fake news"* pandemic
- Social Media – wide use of *"bots"* to manipulate rankings and trends
- Politics – pervasive practice of *"spin"* and *"bullshit"* deceptions
- Commerce – case studies *Enron* and *Volkswagen* scandals
- Finance – money laundering via Crypto-currencies and fake companies
- Research – "*fake*" academic journals and contaminated data sets
- Internet– the spam pandemic, phishing scams etc

- What do these deceptions all have in common?
- *They all exploit the digital information infrastructure*

How Does The Digital Infrastructure Enable a Deception Pandemic?

- Near speed-of-light transmission of data globally
- Very low cost of data transmission / replication / broadcast
- Pervasive reach and footprint in developed nations
- Digital copies are exact copies (if not corrupted)
- Digital fake documents are easy and cheap to produce
- Digital fake documents can be very difficult to unmask as fakes
- Digital fake identities on the Internet are easy and cheap to produce
- Digital fake identities on the Internet can be very difficult to unmask as fakes
- Overcoming limitations of print media removes same obstacles in fakery

Human Susceptibility to Deceptions and Self-Deceptions

- Humans can be very lazy and usually do not check the veracity of data
- Humans too often avoid critical thinking and fall for logical fallacies
- Humans are mostly susceptible to prior cognitive biases
- Humans *self-deceive* to avoid cognitive dissonance (Ramachandran)
- Humans *self-deceive* to better deceive others (Trivers)
- Humans are prone to *self-deceptive Groupthink* (Janis)
- Humans are prone to *Pluralistic Ignorance* seeking conformity
- Humans have an innate curiosity about the improbable (Shannon's information theory states that improbable but ***true*** messages carry more information), which is exploited using improbable and ***untrue*** messages

Janis' Groupthink Problem

- *"Groupthink is a psychological phenomenon that occurs within a group of people, in which the desire for harmony or conformity in the group results in an irrational or dysfunctional decision-making outcome. Group members try to minimize conflict and reach a consensus decision without critical evaluation of alternative viewpoints, by actively suppressing dissenting viewpoints, and by isolating themselves from outside influences."* - Wikipedia definition of Groupthink
- Janis: *"I use the term groupthink as a quick and easy way to refer to the mode of thinking that persons engage in when concurrence-seeking becomes so dominant in a cohesive in-group that it tends to override realistic appraisal of alternative courses of action."*

The Dunning-Kruger Effect Problem

- *"The Dunning–Kruger effect is a cognitive bias in which relatively unskilled persons suffer illusory superiority, mistakenly assessing their ability to be much higher than it really is. Dunning and Kruger attributed this bias to a metacognitive inability of the unskilled to recognize their own ineptitude and evaluate their own ability accurately. Their research also suggests corollaries: highly skilled individuals may underestimate their relative competence and may erroneously assume that tasks which are easy for them are also easy for others."* – Wikipedia definition of Dunning-Kruger Effect
- RationalWiki: *"**people who are too stupid to know how stupid they are"***

Nichol's "Death of Expertise" Problem

- Nichols: *"These are dangerous times. Never have so many people had access to so much knowledge, and yet been so resistant to learning anything." "People are now exposed to more information than ever before, provided both by technology and by increasing access to every level of education. These societal gains, however, have also helped fuel a surge in narcissistic and misguided intellectual egalitarianism that has crippled informed debates on any number of issues. Today, everyone knows everything: with only a quick trip through WebMD or Wikipedia, average citizens believe themselves to be on an equal intellectual footing with doctors and diplomats. All voices, even the most ridiculous, demand to be taken with equal seriousness, and any claim to the contrary is dismissed as undemocratic elitism."* – Amazon.com

Frankfurt's "*Bullshit*" Problem

- *"**On Bullshit** (2005), by philosopher Harry G. Frankfurt, is an essay that presents a theory of bullshit that defines the concept and analyzes the applications of bullshit in the contexts of communication. Frankfurt determines that bullshit is speech intended to persuade (a.k.a. rhetoric), without regard for truth. The liar cares about the truth and attempts to hide it; the bullshitter doesn't care if what they say is true or false, but rather only cares whether or not their listener is persuaded."*
- BS has become a pervasive and destructive practice in politics, sales and commerce, mass media, social media and other domains.
- The effort to refute BS usually greatly exceeds the effort to produce it.
- *Where BS becomes normalised in a culture, it is difficult to stamp out!*

A “Post Truth World”?

- *“**Post-truth politics** (also called **post-factual politics**[1] and **post-reality politics**[2]) is a political culture in which debate is framed largely by appeals to emotion disconnected from the details of policy, and by the repeated assertion of talking points to which factual rebuttals are ignored. Post-truth differs from traditional contesting and falsifying of facts by relegating facts and expert opinions to be of secondary importance relative to appeal to emotion. While this has been described as a contemporary problem, some observers have described it as a long-standing part of political life that was less notable before the advent of the Internet and related social changes.”*
  - Wikipedia

Are We Seeing a “Perfect Storm” of Mis-Information and Dis-Information?

- We are observing a confluence of technological advancements in digital technology and cultural shifts away from respect for facts and truth, and an arbitrary acceptance of nonsense where it appeals to the cognitive biases or other agendas of the audience;
- This produces a fertile environment for deceivers promoting self-serving and other agendas;
- Traditional deception methods are being adapted and enhanced by digital technology;
- New techniques such as social media “*Bots*”, sowing confusion *en-masse*, and *“flooding / saturation”* attacks now increasing observed in digital media, used to influence politics and public debate.

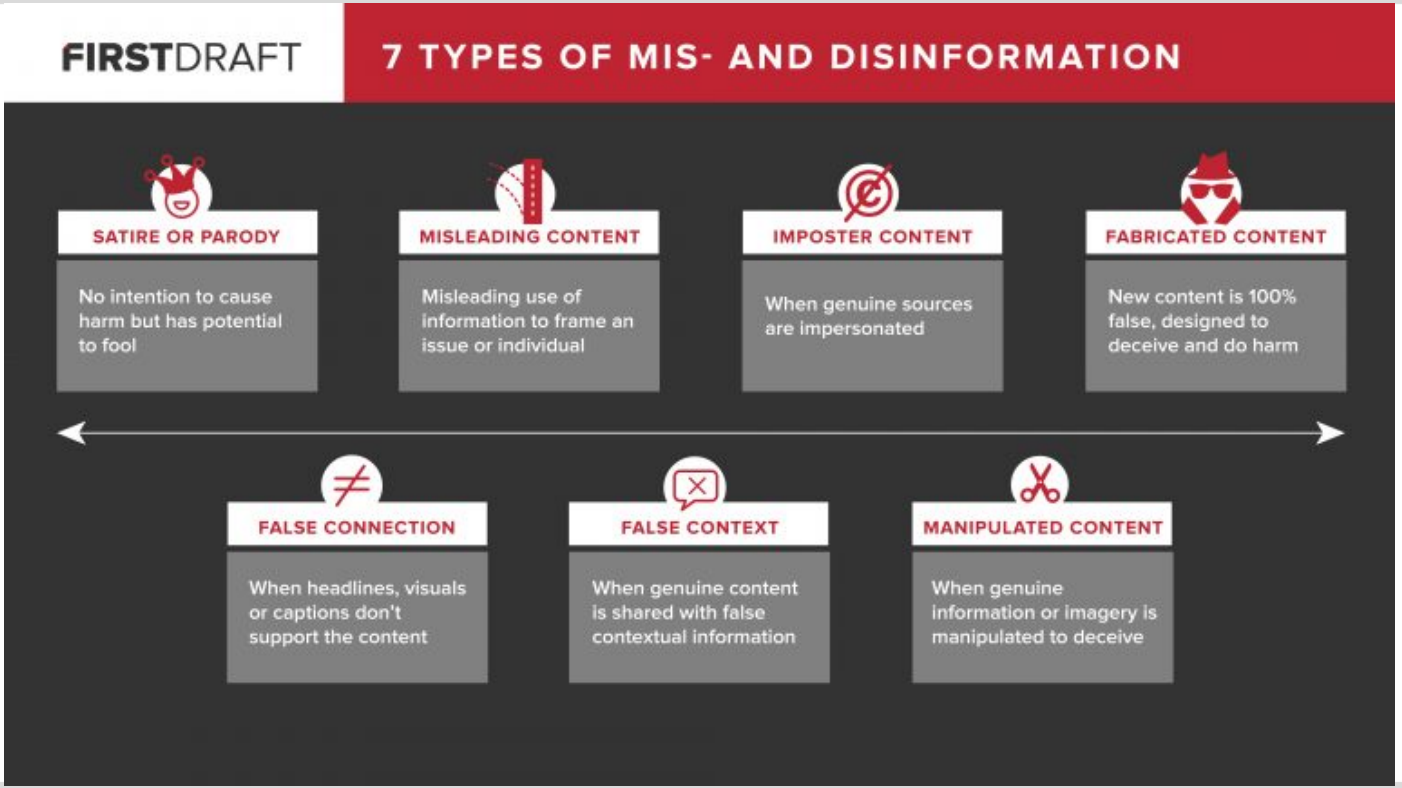
MONASH University

The *"Fake News"* Problem

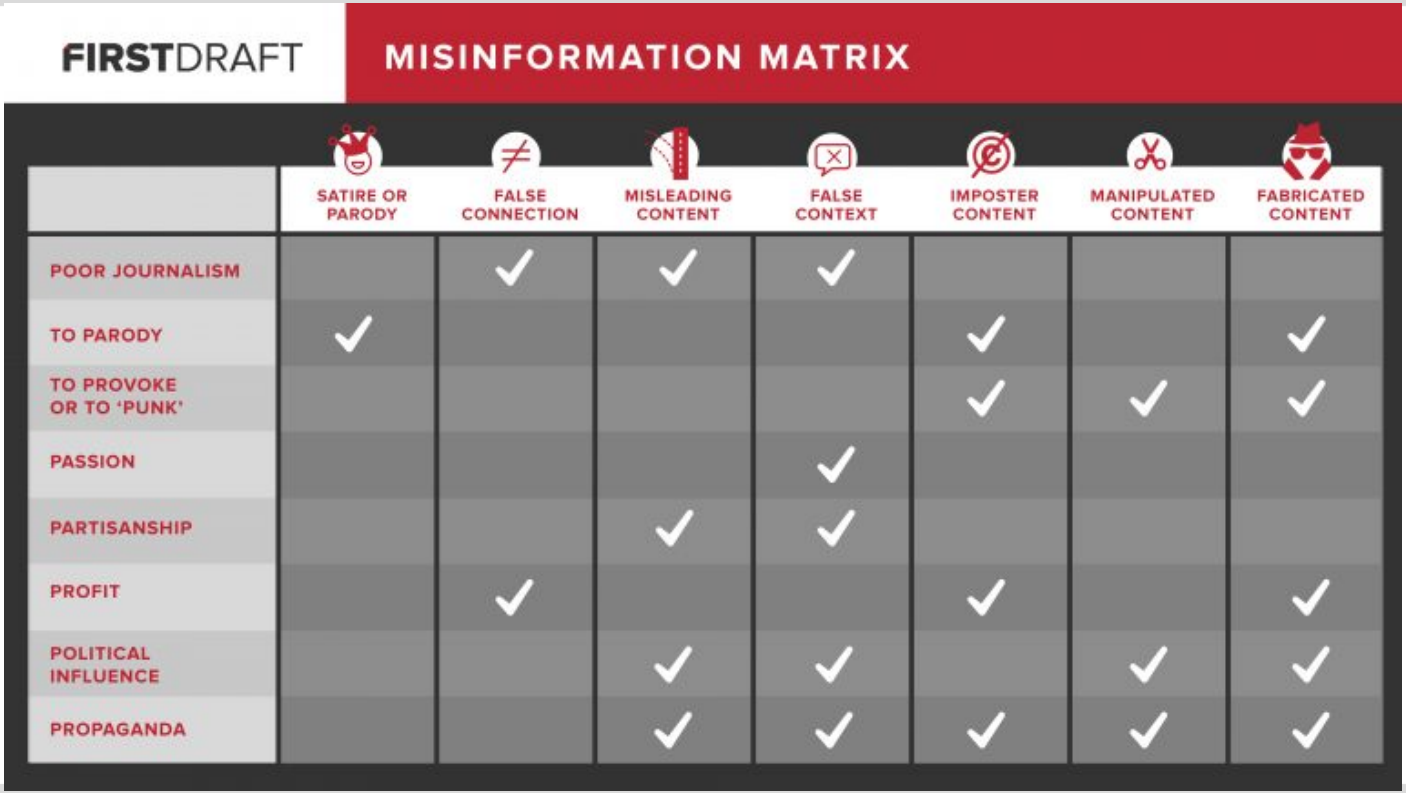
What is “Fake News”?

- Lazer et al: *“fabricated information that mimics news media content in form but not in organizational process or intent.” “Fake news overlaps with other information disorders, such as misinformation (false or misleading information) and disinformation (false information that is purposely spread to deceive people).”*
- Campan et al: *“clickbait, propaganda, commentary/opinion and humour/satire”*, mis-information where veracity is unknown, dis-information where there is intent to deceive.
- Wardle mapped seven means and eight motives for the production and distribution of misinformation, based on the empirical observation of social and mass media “fake news”.

Mis/Disinformation – Wardle 2017



Wardle's Misinformation Matrix – Wardle 2017



	SATIRE OR PARODY	FALSE CONNECTION	MISLEADING CONTENT	FALSE CONTEXT	IMPOSTER CONTENT	MANIPULATED CONTENT	FABRICATED CONTENT
POOR JOURNALISM		✓	✓	✓			
TO PARODY	✓				✓		✓
TO PROVOKE OR TO 'PUNK'					✓	✓	✓
PASSION				✓			
PARTISANSHIP			✓	✓			
PROFIT		✓			✓		✓
POLITICAL INFLUENCE			✓	✓		✓	✓
PROPAGANDA			✓	✓	✓	✓	✓

MONASH University

How Does Deception Work?

Established Study of Deceptions

- Deceptions in social systems have been studied by strategy, social sciences and humanities scholars for decades, with a focus on the taxonomical classification of deceptions, and collection of case studies;
- Notable works by Haswell, Heuer, Fleming and Zyglidopoulos, Bell and Whaley, Flynn et al, Pomerantsev and others;
- Psychology researchers are more recently focusing on deceptions, with a wealth of recent papers – NB Berne's 1966 early study of social games;
- Hard sciences research on deceptions is much less well developed.
- Game theorists have been hampered by a lack of a well established mathematical theory, as extant information theory models for deception are very recent and not widely known in the science community.

Classical Intelligence Deception Techniques (Haswell)

- **The Lure** – this technique presents the opponent with a sudden advantage they may exploit.
- **The Repetitive Process** – this technique conditions the opponent by repetition to accept harmless behaviour that is used as a cover for subsequent operations.
- **The Unintentional Mistake** – this technique leads an opponent to believe that valuable information has come into his hands by mistake, for instance by negligence or incompetence.
- **The Obvious Solution** – this technique provides deceptive information to support the idea that the obvious method will be used, while hiding information related to the actual method.
- **The Piece of Bad Luck** – this technique is similar to the Unintentional Mistake, except the bad luck cannot be attributed to anyone.

Political and Commercial Deceptions

- **Deception By Omission** - The attacker hides information which would be unhelpful or deleterious in driving the victim of the deception to a specific misperception of reality.
- **Deception By Saturation** (Kopp) / **Flooding** (Libicki) – The attacker will inundate the victim with messages, most of which are redundant or irrelevant, with the aim of saturating the victim's channel so the victim cannot gather information which might contradict the attacker's message.
- **Deception By Spin** (Bernays and Goebbels) - the attacker presents an unpalatable yet accepted fact, but encourages the victim to assess that fact from a perspective which is less damaging to the attacker.

Bernays' Deception by "Spin"

- A *Spin Attack* is based on the idea of presenting an unpalatable or other acknowledged or accepted fact, but encouraging the victim to assess that fact from a perspective which is less damaging to the attacker.
- *Indirect Spin Attacks* attempt to conceal the connection between the unwanted fact and the *Spin Attack*.
- Trivial example of the basic form might be thus – "*here is an fact which is true, but it isn't really that bad because of the following circumstances* ….".
- The explanation of *'following circumstances*' compels or encourages the victim to devalue the unwanted consequences of the unpalatable fact.
- The attacker presents *'following circumstances'* which may in themselves not be untruthful, but achieve a deceptive aim by altering the victim's interpretation of the message to the advantage of the attacker.

More on Spin Attacks

- Spin attacks, like *Deception By Omission* attacks, rely on the victim having little or no *a priori* knowledge or understanding, and the victim not being prepared to critically analyse a statement by the attacker.
- The use of spin attacks thus often relies on the trust of the victim, or victims who are fearful of losing confidence in the attacker.
- Spin Attacks are popular since if well executed, the attacker need not make obviously false statements to achieve the deceptive aim.
- As the victim uses its own internal processing resources to infer false conclusions from the received message, the victim has been effectively subverted to an internal state which is intended by the attacker (Brumley et al, 2006).

Mass Media *"Perception Management"* Techniques

- Notable examples are:
  1. Germany's Third Reich
  2. Soviet Union, Warsaw Pact, Russian Federation, DPRK and PRC
  3. Al Qaeda, Iran and affiliated Islamo-fascist movements
- Specific pattern of technique and the use of a sustained and internally consistent long term deception campaign, characteristically targeted at followers of the regime or movement.
- More than often *"perception management"* techniques intended to attack external opponents of such regimes are unique, and indeed different from those targeting the captive population.

Proxy Delivery of Deceptions: Mass Media Attributes

- Focussed on the delivery of *"infotainment"* rather than dedicated news and news analysis.
- Timeliness of delivery has precedence over the depth of analysis or accuracy of the material.
- By-product of a commercial market dynamic - competing media players must attract the interest of viewers to achieve favourable ratings and thus attract subscriptions or advertising revenues (*de facto "clickbait"*).
- Commercial application of Goebbels' dictum that *"propaganda must be entertaining"* (Goebbels, 1943).
- The implicit aim of this propaganda is transmission of the message that *"this media organisation is more attractive than its competitors"*.

Proxy Delivery of Deceptions: Observable Realities

- Viewers and readers are most attracted to footage or stories which are dramatic, violent or involve intense controversy.
- Media organisations aim to appeal to existing prejudices or preconceptions on the part of the audience (cognitive bias).
  - i.e. the same mechanism observed in propaganda distribution, as presentation of materials which challenge audience prejudices or preconceptions will be less likely to be received favourably
- Media organisation audience deception is a compound strategy centred on audience interest and *a priori* prejudices and aimed at maximising audience visitation rates at the expense of competitors.
- *Competitive game of "who has the best honeypot?" with the game payoff in the frequency of visitation and thus traffic related revenue.*

Proxy Delivery of Deceptions: Propaganda vs. Western Democracies

- Attacker must wrap deceptive message in an envelope of material which is attractive to global media organisations or social media users.
- *Deceptive message must provide content which is dramatic, violent, intensely controversial, or any combination of the three, and which appeals to audience prejudices where possible.*
- In the context of Information Warfare, distribution of deceptive propaganda using the global mass media as a conduit employs compound strategies combining a range of methods.
- *Destruction of the delivery channel is usually avoided since it compromises the intermediate aim of the strategy, which is exploitation of the delivery channel.*

Observations

- There is a large body of extant research focused on deceptions
- The basic methods and techniques employed in intelligence deceptions, nation state propaganda deceptions, commercial / sales deceptions and political deceptions differ only in detail
- Proxy delivery of deceptions via media and social media is a growing trend – using others to deceive on your behalf yields self-propagating and self-funding deceptions thus minimising costs to the deceiver
- Despite these established deceptions being well understood by scholars of deception, they are usually very poorly understood by politicians, mass media, and most of the public

MONASH University

Information Theoretic Modeling of Deception

Background to Information Theoretic Modeling of Deceptions

- The four information theoretic models of deception were identified almost concurrently by Colonel Andrew Borden, PhD, USAF, in the US, and Carlo Kopp, at Monash University CSSE, in 1999.
- Dr Borden published two months before Kopp, in APJ *Air Chronicles*, the United States Air Force professional journal. Kopp published in the Australian industry journal *Systems*, formerly *Australian Unix User's Review*.
- Borden's model does not include the "*subversion"* model as a defined model, and follows the US DoD convention of opaquely conflating it within the "*denial"* model.
- The *subversion* strategy was first published by Kopp, and credit for its initial identification must go to the late Prof C.S. Wallace, foundation Chair of Computer Science at Monash University.

Observation

- Claude Shannon developed mathematical information theory at Bell Labs during the mid 1940s.
- It has since then provided the theoretical foundations for digital communications and digital data storage.
- Of the four observed types of deception that arise in social and biological systems, three can be easily described by manipulations of Shannon's channel capacity theorem, the fourth by Turing's work.
- The Borden-Kopp model has been cited in numerous information warfare publications, and is the subject of Ch.4 in Poisel's 2013 textbook: Refer Poisel R.A. *Information Warfare and Electronic Warfare Systems*. Norwood, Massachussetts: Artech House; 2013.

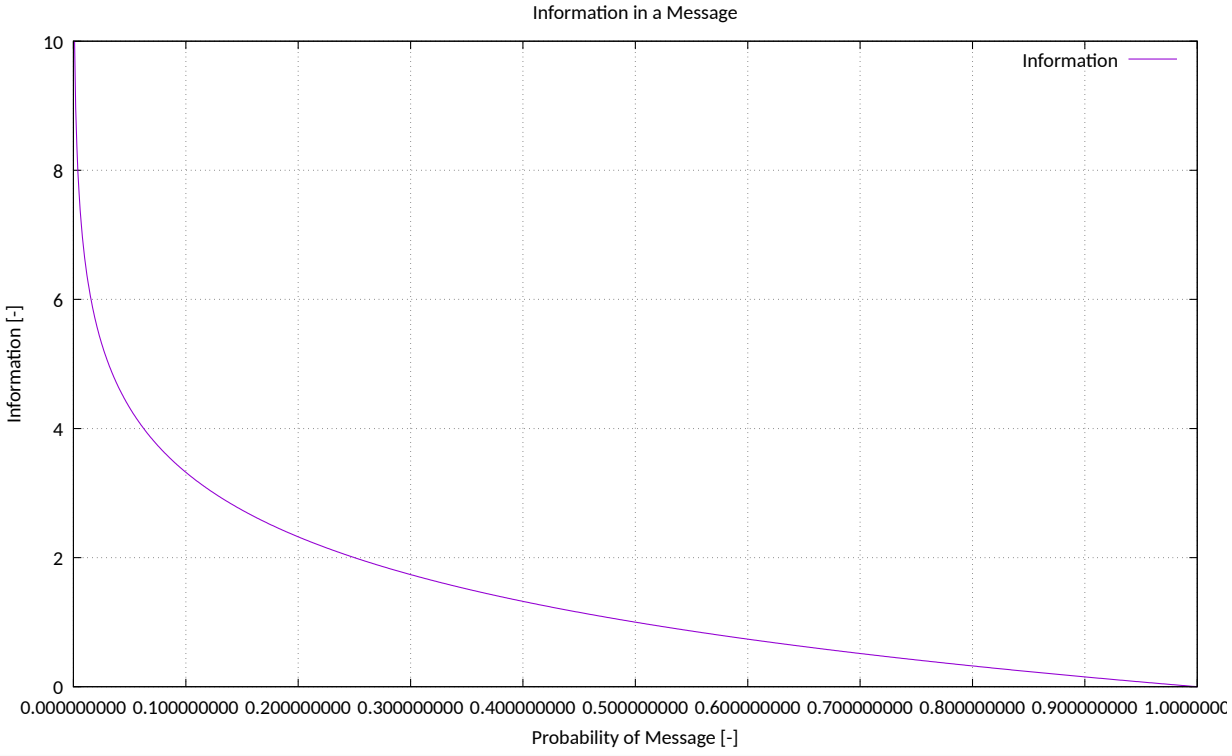
Information in a Message

- The Entropy Theorem presents the total entropy (information) in the system, across all of the possible messages in the system.
- For many problems we are interested in the information in one of the N messages. This can be represented thus:

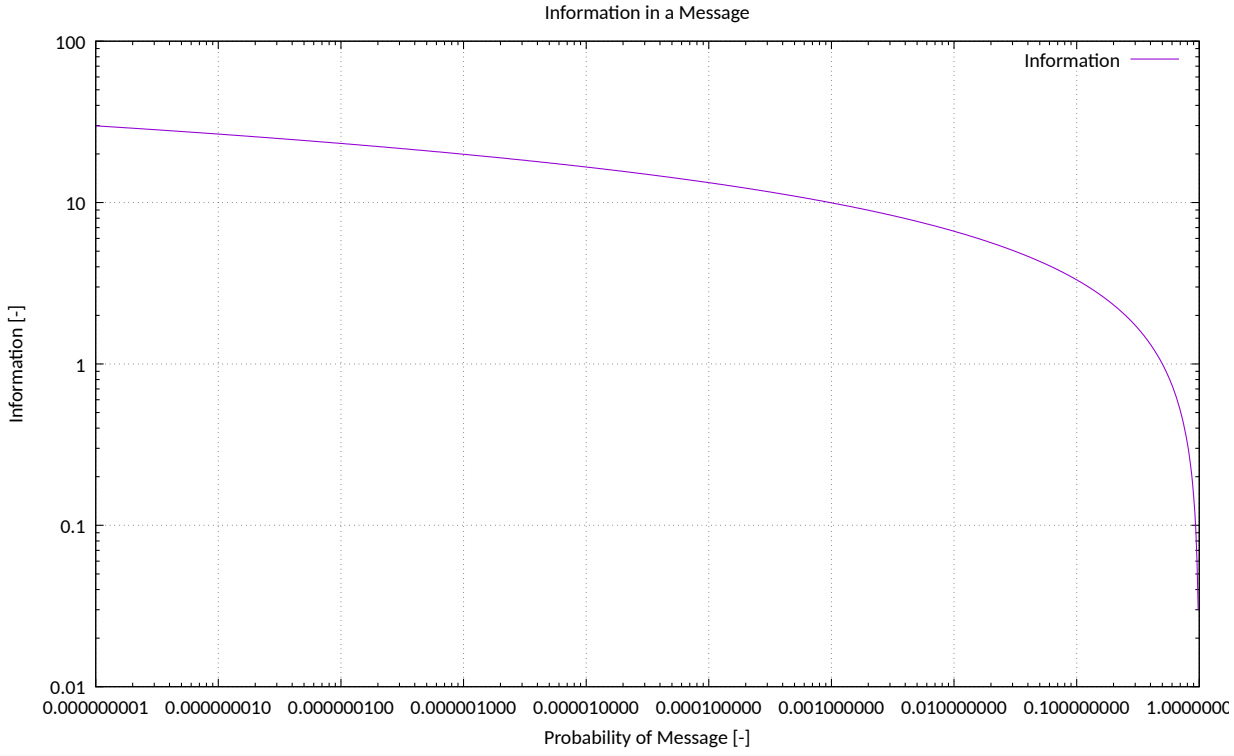
$$I(m) = -\log_2(p(m))$$

- As is evident, highly probable messages contain little information and vice versa.

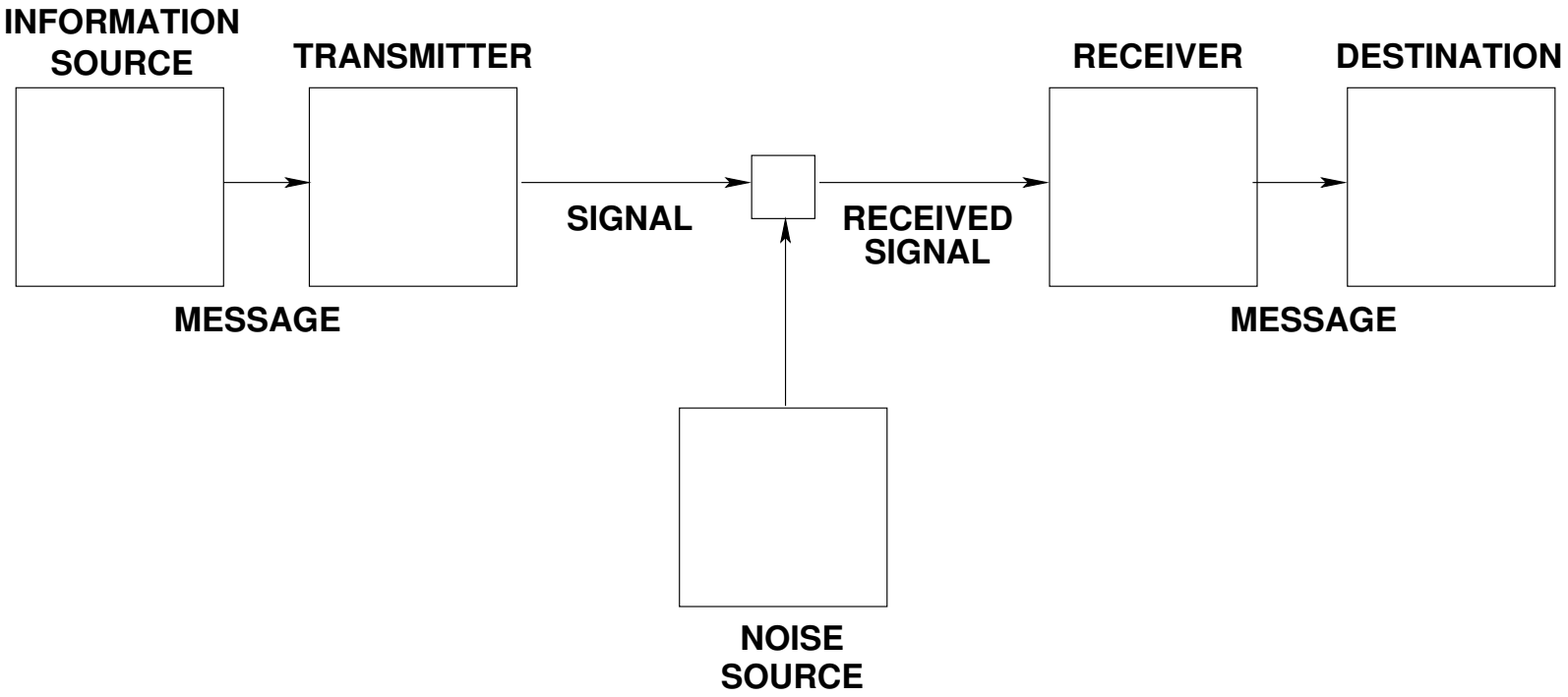
Information vs Likelihood of Message



Information vs Likelihood of Message



Shannon's Channel Capacity Model



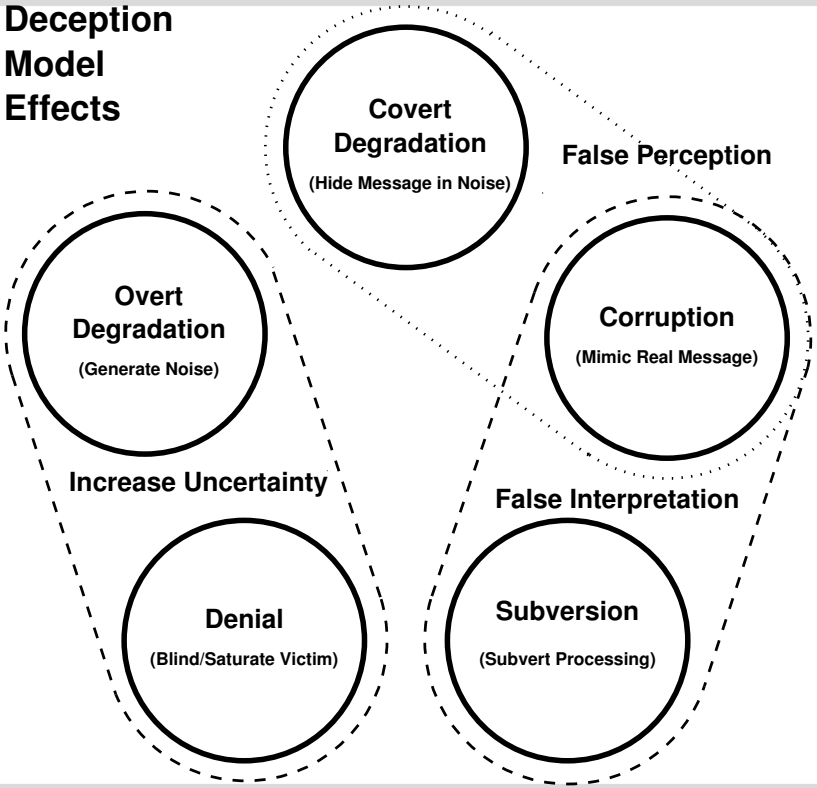
The Four Canonical Models of Deception

- **Degradation** amounts to making the signal sufficiently noise-like, that a receiver cannot discern its presence from that of the noise in the channel, or burying it in generated noise.
- **Corruption** amounts to mimicking a known message so well, that a receiver cannot distinguish the fake message from the real message .
- **Denial** amounts to injecting so much noise into the channel, that the receiver cannot decode the message; alternately the outright destruction of the receiver subsystem.
- **Subversion** amounts to altering the algorithms or other parameters used by the victim to make decisions or take actions; the diversion of the thread of execution within a Turing machine, which maps on to the functional behaviour of the victim system.

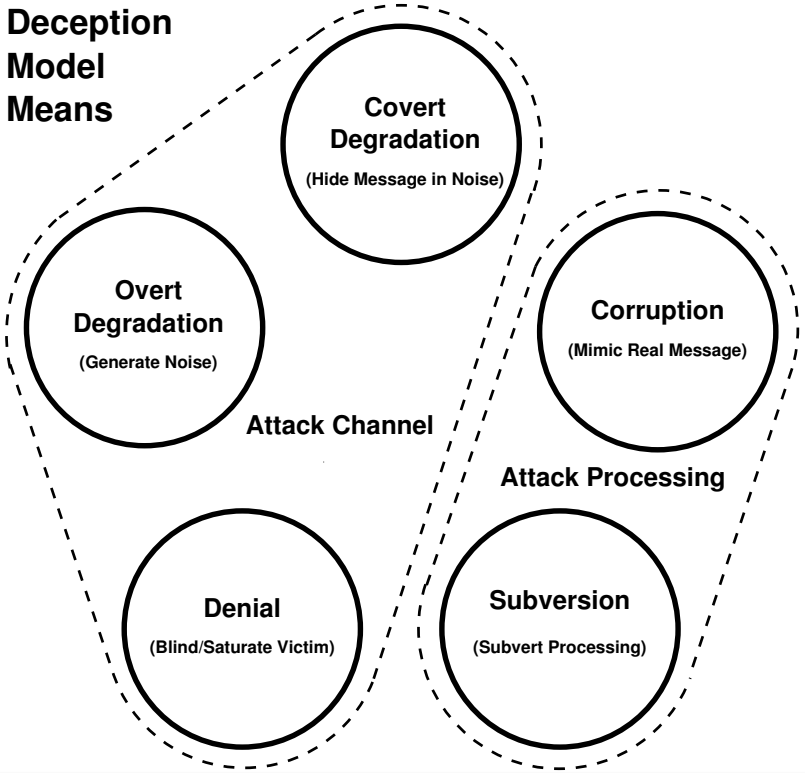
The Effects of the Four Canonical Models of Deception

- **Degradation:** introduces uncertainty (overt) or a false belief (covert) in the victim's perception of its environment.
- **Corruption:** introduces a false belief in the victim's perception of its environment.
- **Denial:** introduces uncertainty in the victim's perception of its environment.
- **Subversion:** alters the manner in which the victim interprets its beliefs about its environment, or may even alter the victim's agendas and motivations.

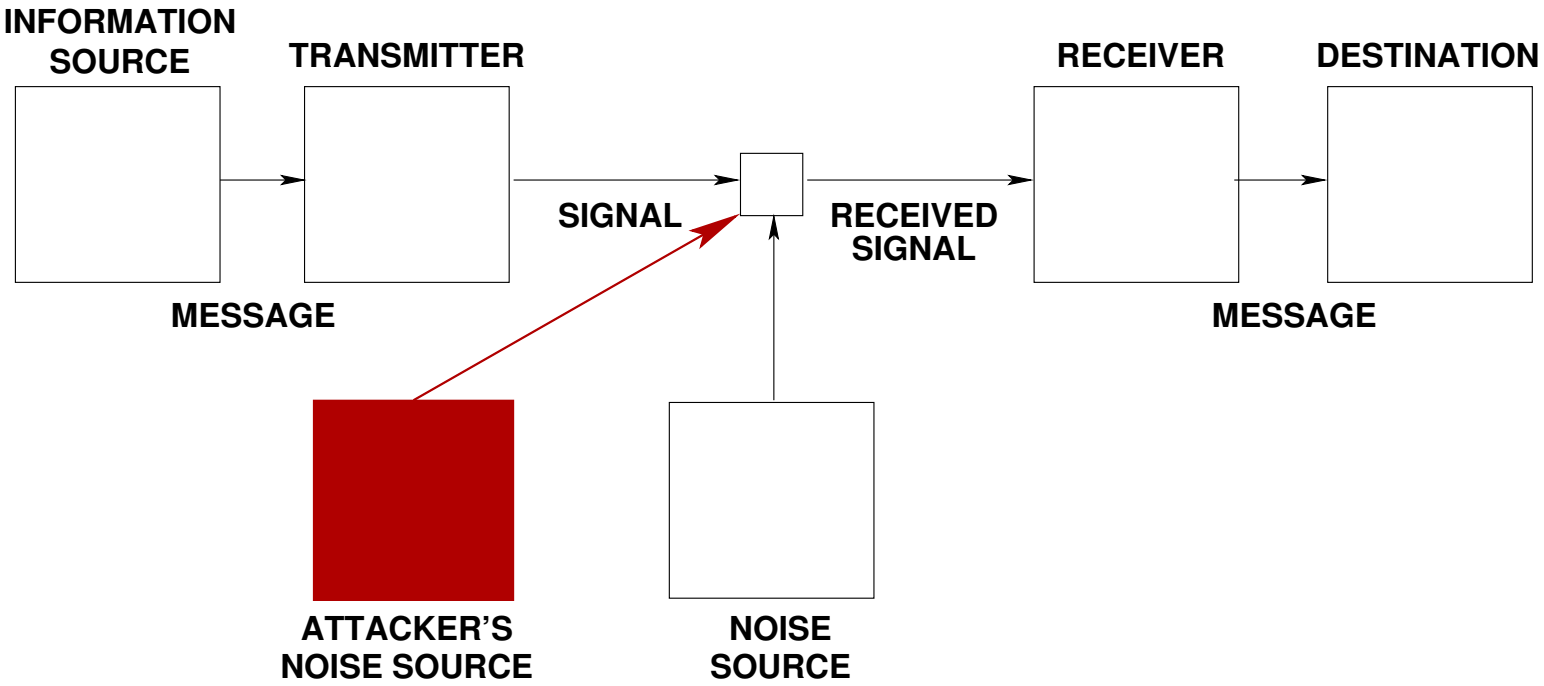
Deception Model Effects



Deception Model Means

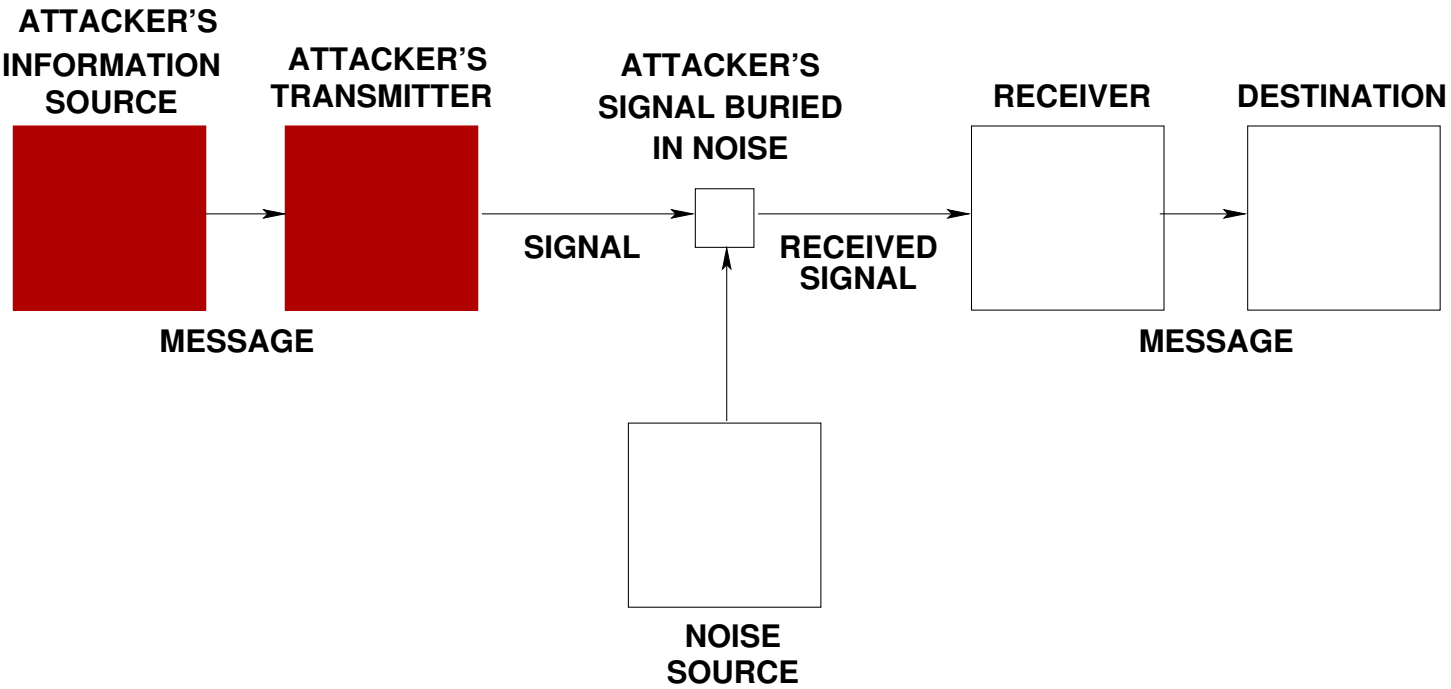


Degradation Model – Overt / Active Form



Degradation Deception Model

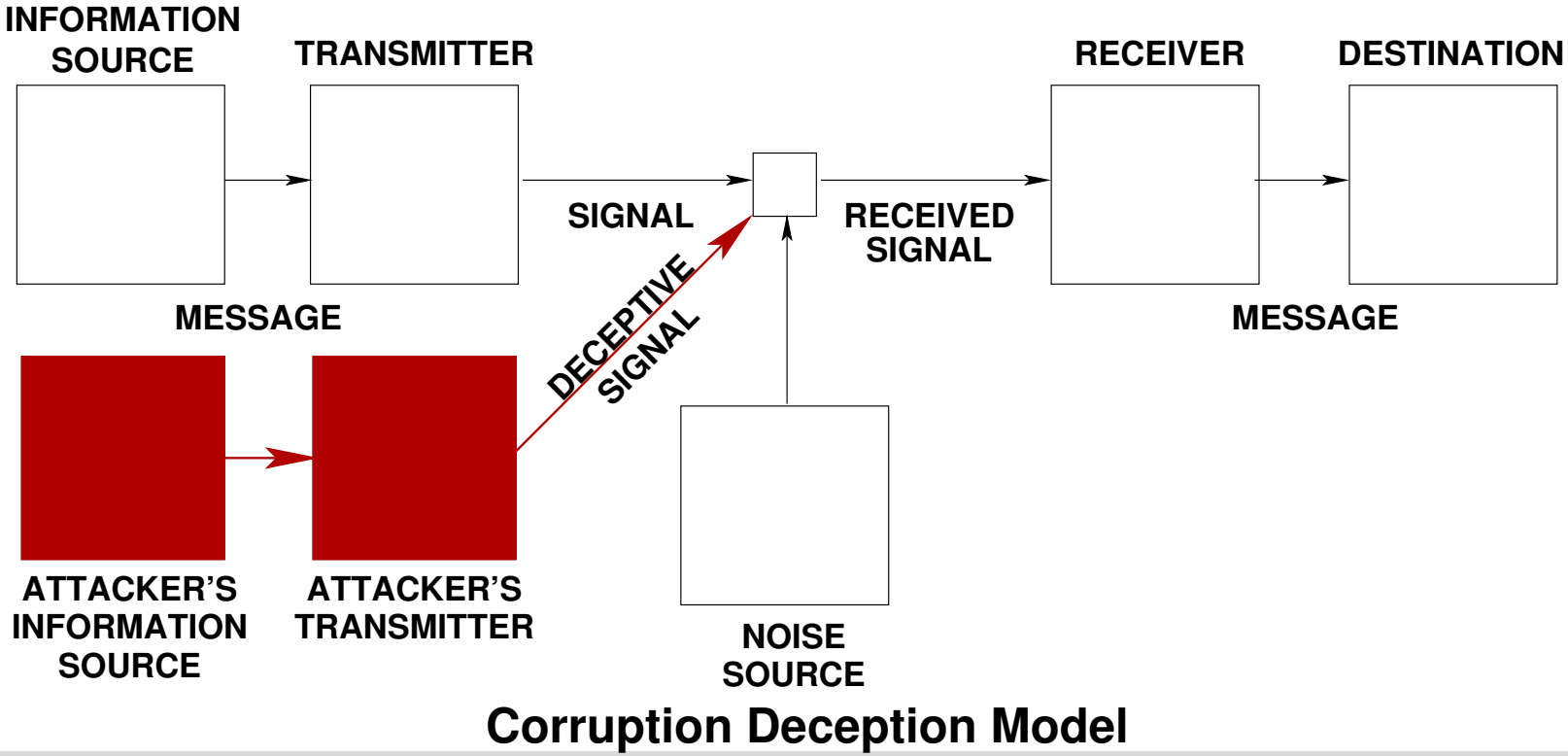
Degradation Model – Covert / Passive Form



Degradation Covert / Passive Deception Model

Example - Degradation

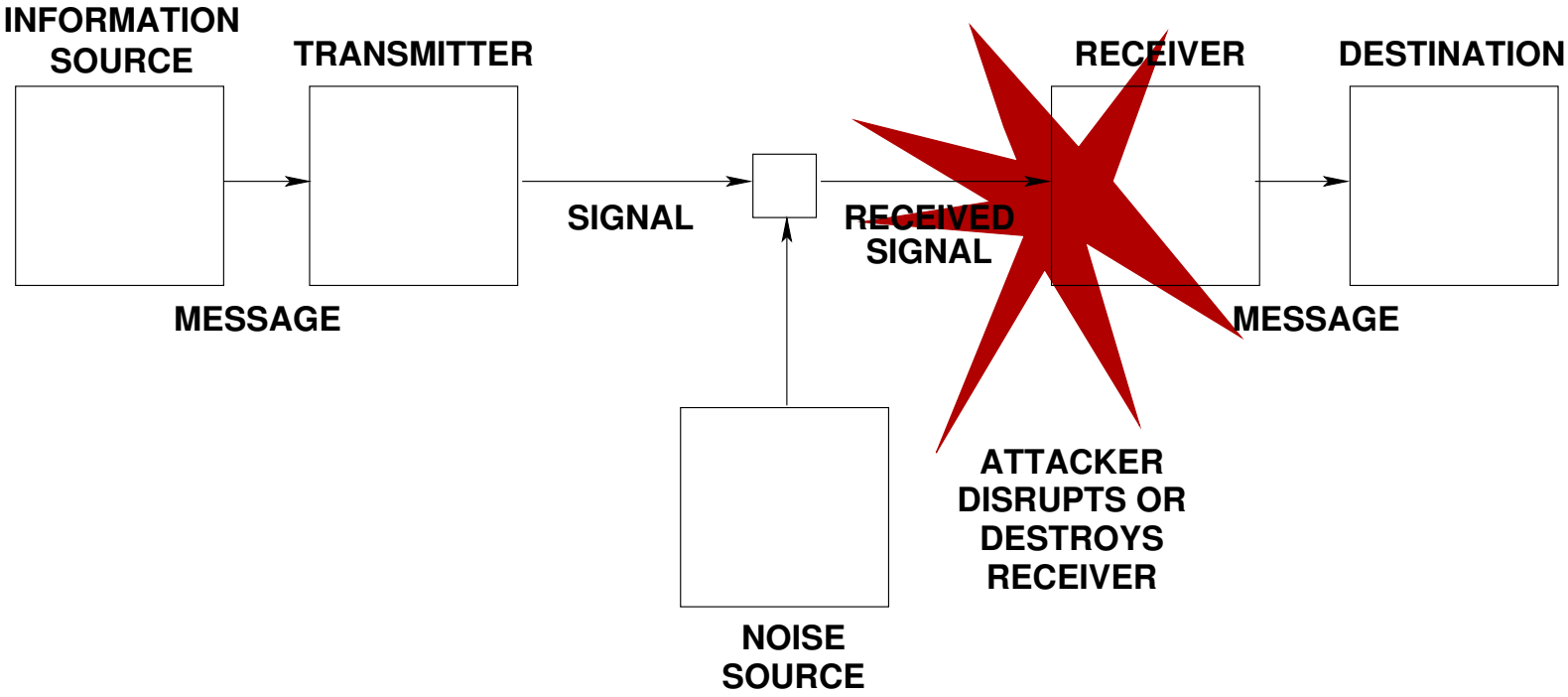
Corruption Deception Model



Corruption Deception Model

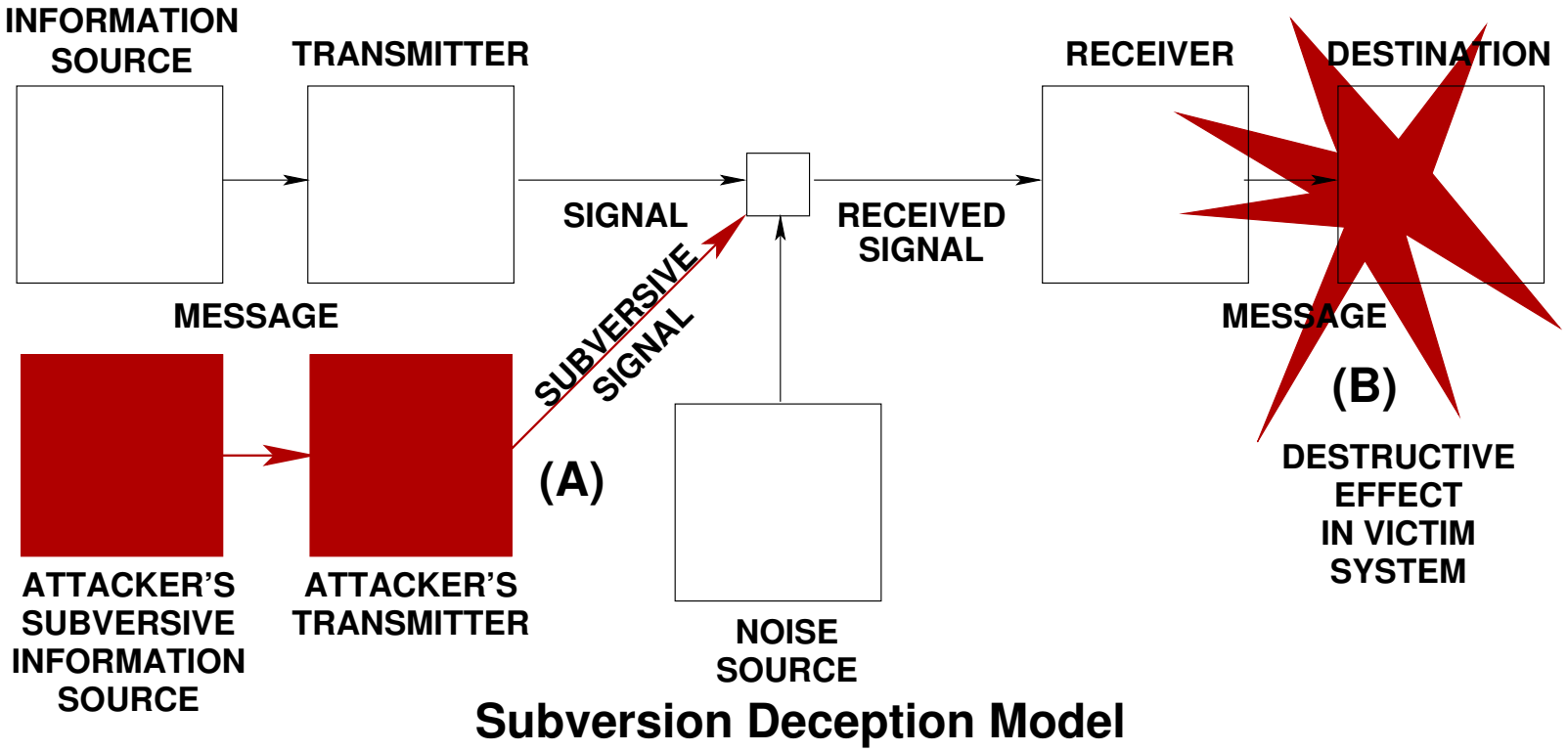
Example - Corruption

Denial Deception Model

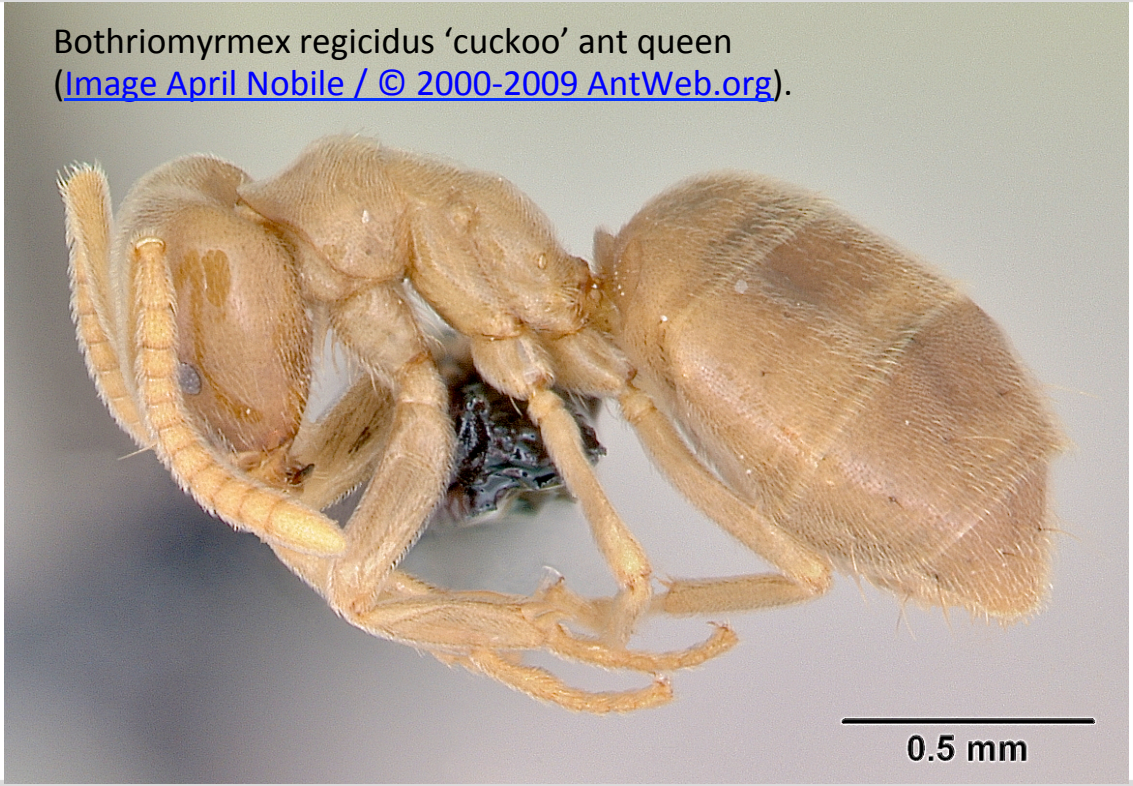


Example – Denial

Subversion Deception Model

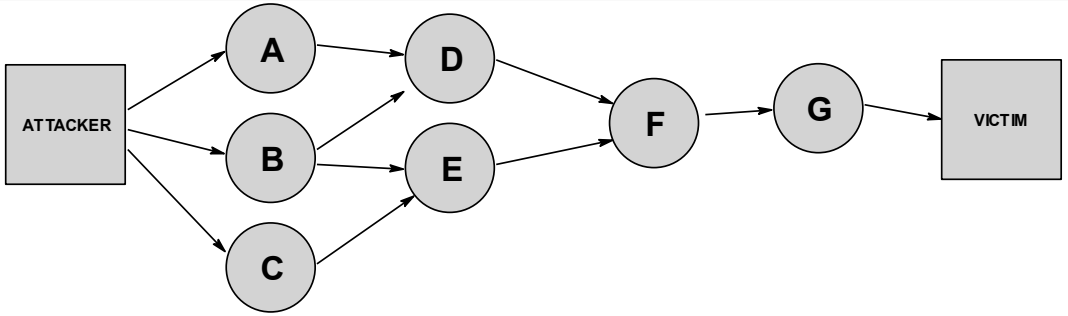


Examples - Subversion

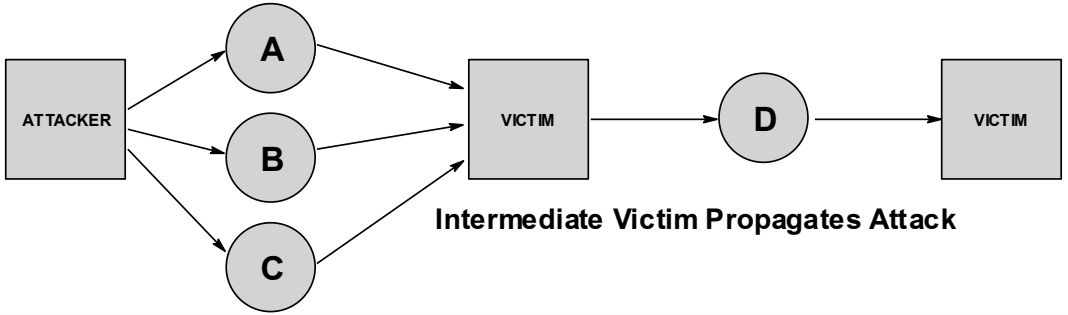


Compound Deceptions

- Often complicated meshes or webs of interconnected deceptions;
- Chained compound attacks are important as a victim becomes a proxy for the attacker, propagating the message to new victims.



A Compound IW Strategy



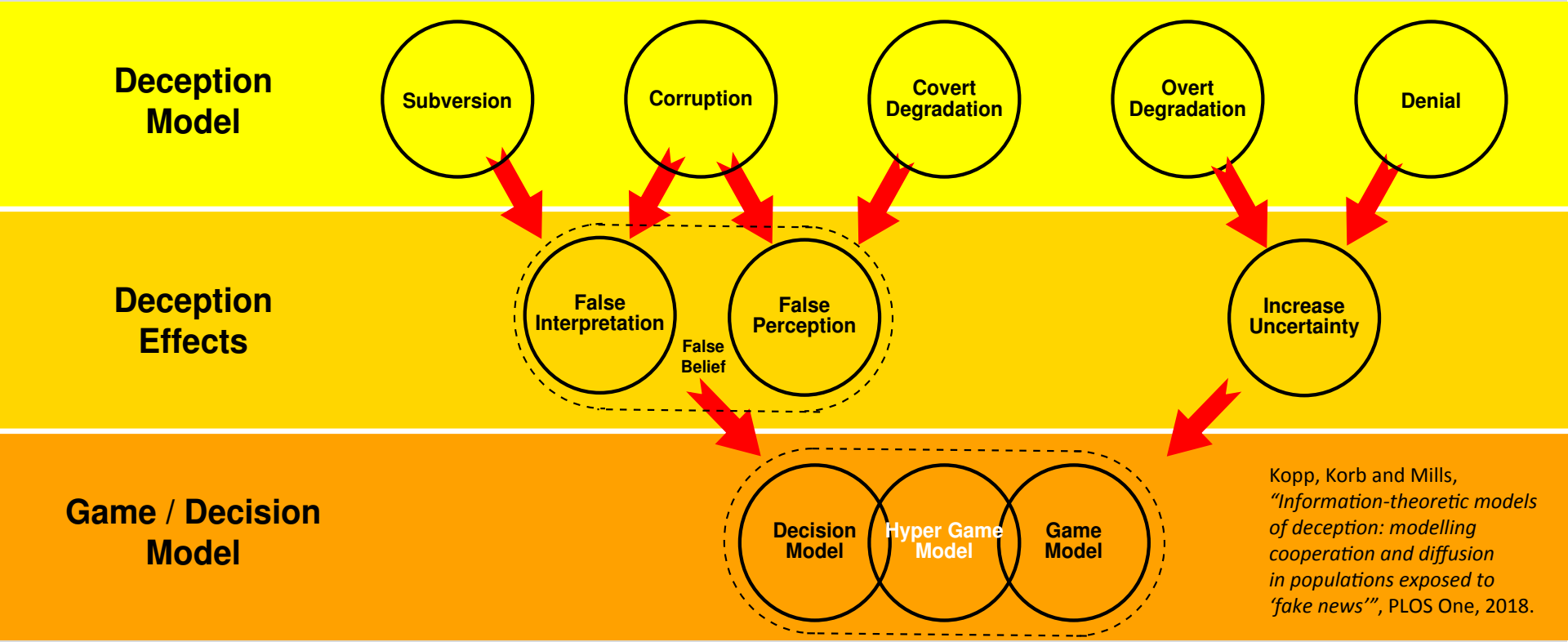
Intermediate Victim Propagates Attack

A Chained Compound IW Strategy

Advances in Information Theoretical Modelling of Deception (2018)

- There has been a long standing unsolved problem in how should deception be incorporated into *Game and Decision Theory* models?
- *Information Theory* models produce deception effects as an output;
- *Game and Decision Theory* models employ deception effects as an input;
- The resolution of this problem is thus simple – *Information Theory* models are chained with *Game and Decision Theory* models, where the output of the former becomes the input to the latter;
- This approach provides a very simple and clean mapping that allows use of past *Game Theory* constructs by Li and Cruz, and *Decision Theory* constructs by Greenberg;

Integrating Borden-Kopp Model for Deception with Game/Decision Theory



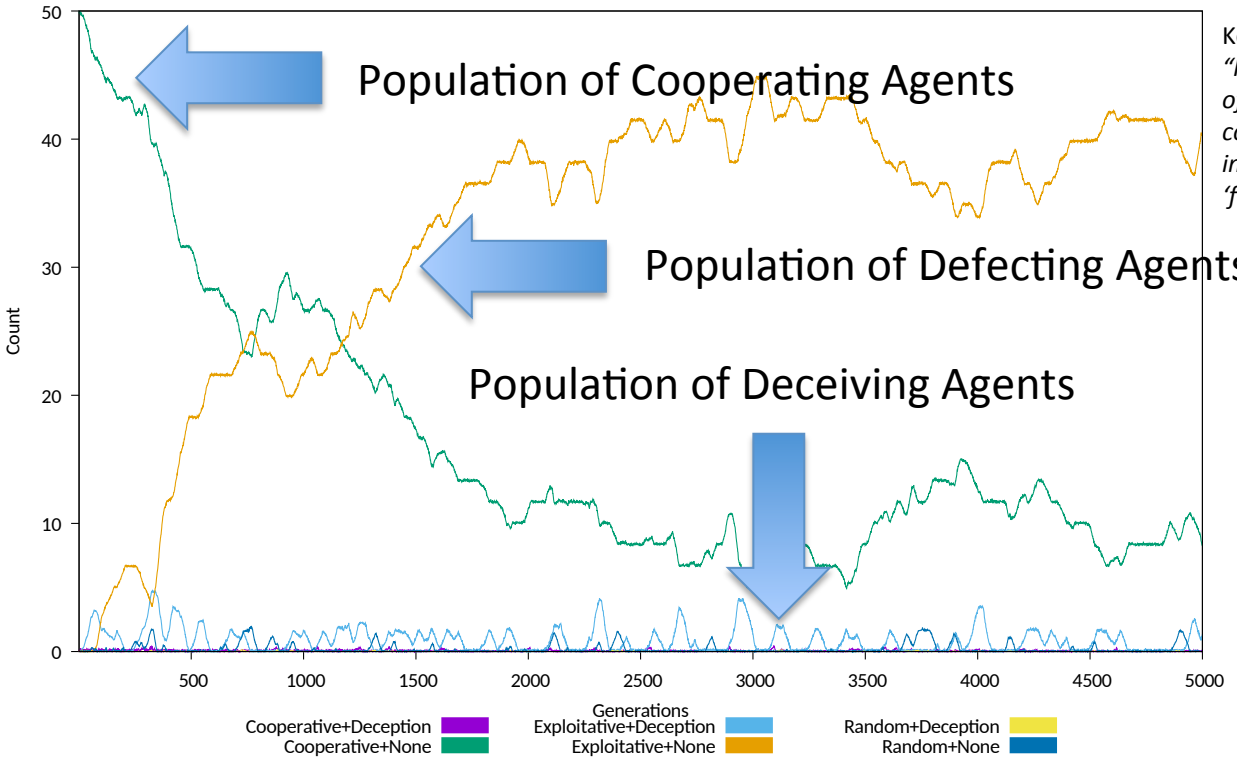
Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

The 2016-18 *"Fake News"* Experiments (Monash Uni FIT Study)

- Intended to explore the effects of *Degradation* and *Corruption* deceptions in agent populations playing randomly paired *Iterated Prisoner's Dilemma* game, using an evolutionary simulation;
- The *Degradation* experiment explored the impact of *Degradation* on consensus forming behaviours in the agent population;
- The *Corruption* experiment explored the diffusion of *Corruption* in the agent population;
- Both experiments were parametrised by the *Cost* of the deception, that reduced fitness of deceiving agents when deceptions failed;
- The results of the experiments, despite the simplicity of the simulation, showed good agreement with empirical observations of social media;

Iterated Prisoner's Dilemma Population with Deceiving Players (Cost 0.05)

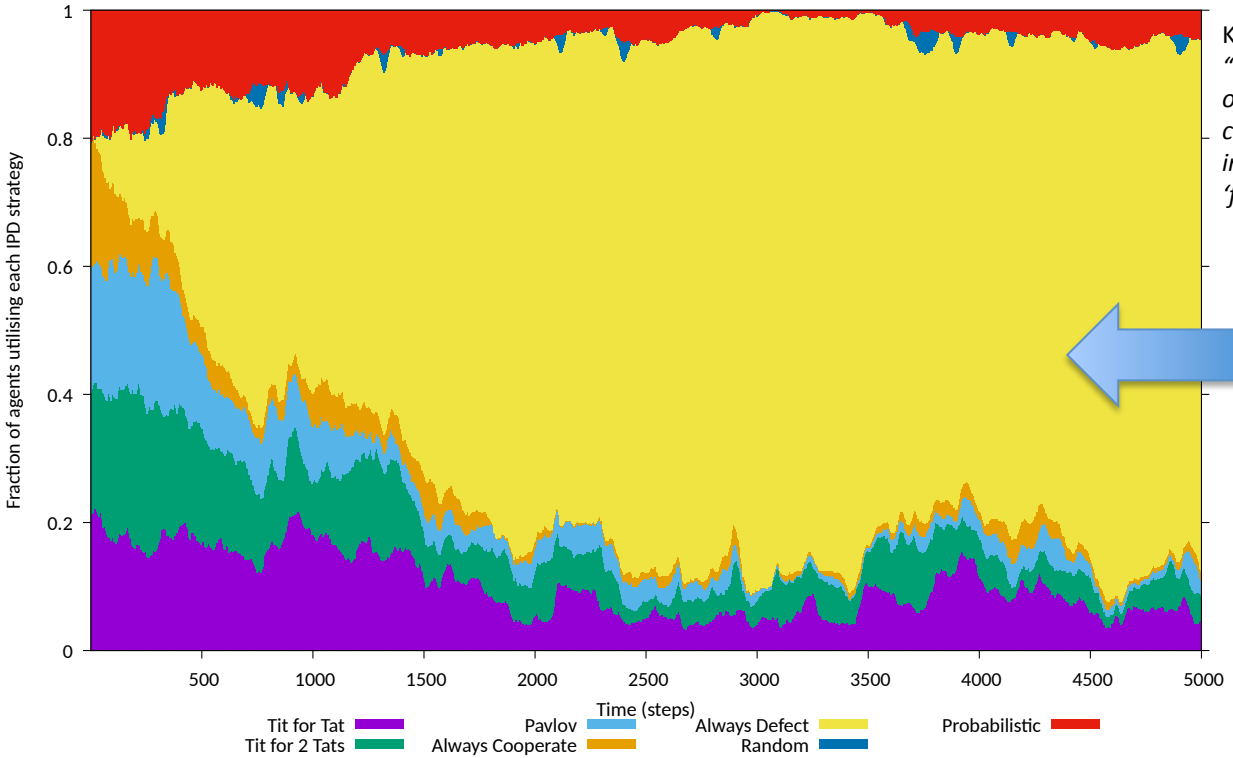
Degradation
Deception
Introducing
Uncertainty
And Disrupting
Cooperation



Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

Iterated Prisoner's Dilemma Population with Deceiving Players (Cost 0.05)

Degradation
Deception
Introducing
Uncertainty
And Disrupting
Cooperation

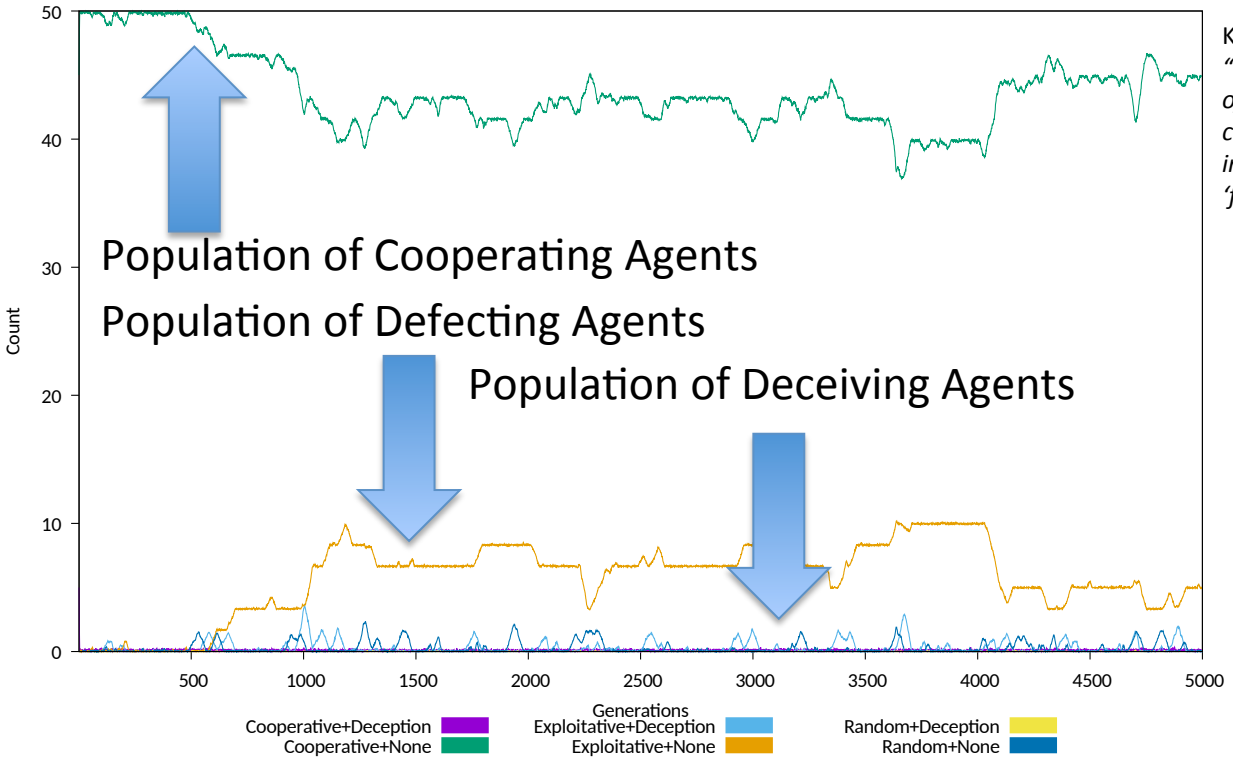


Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

Population of Defecting Agents

Iterated Prisoner's Dilemma Population with Deceiving Players (Cost 0.3)

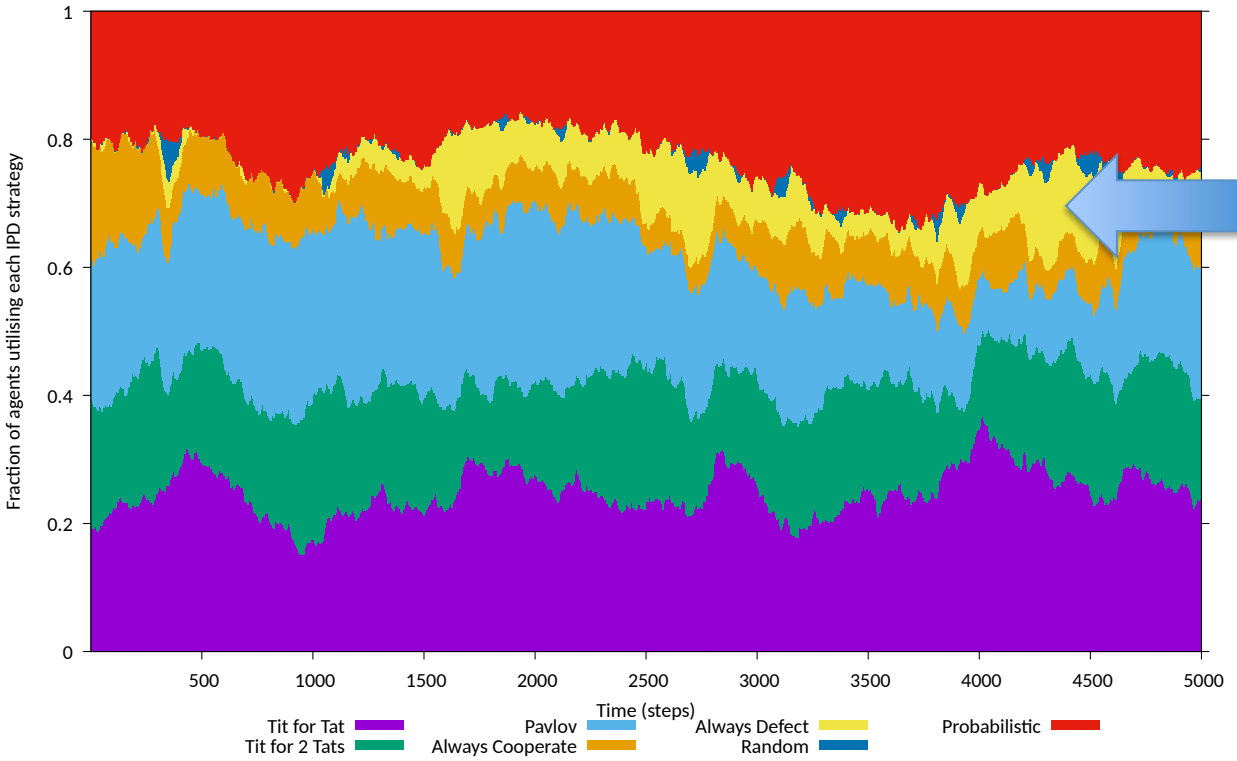
Degradation
Deception
Introducing
Uncertainty
And Disrupting
Cooperation



Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

Iterated Prisoner's Dilemma Population with Deceiving Players (Cost 0.3)

Degradation
Deception
Introducing
Uncertainty
And Disrupting
Cooperation

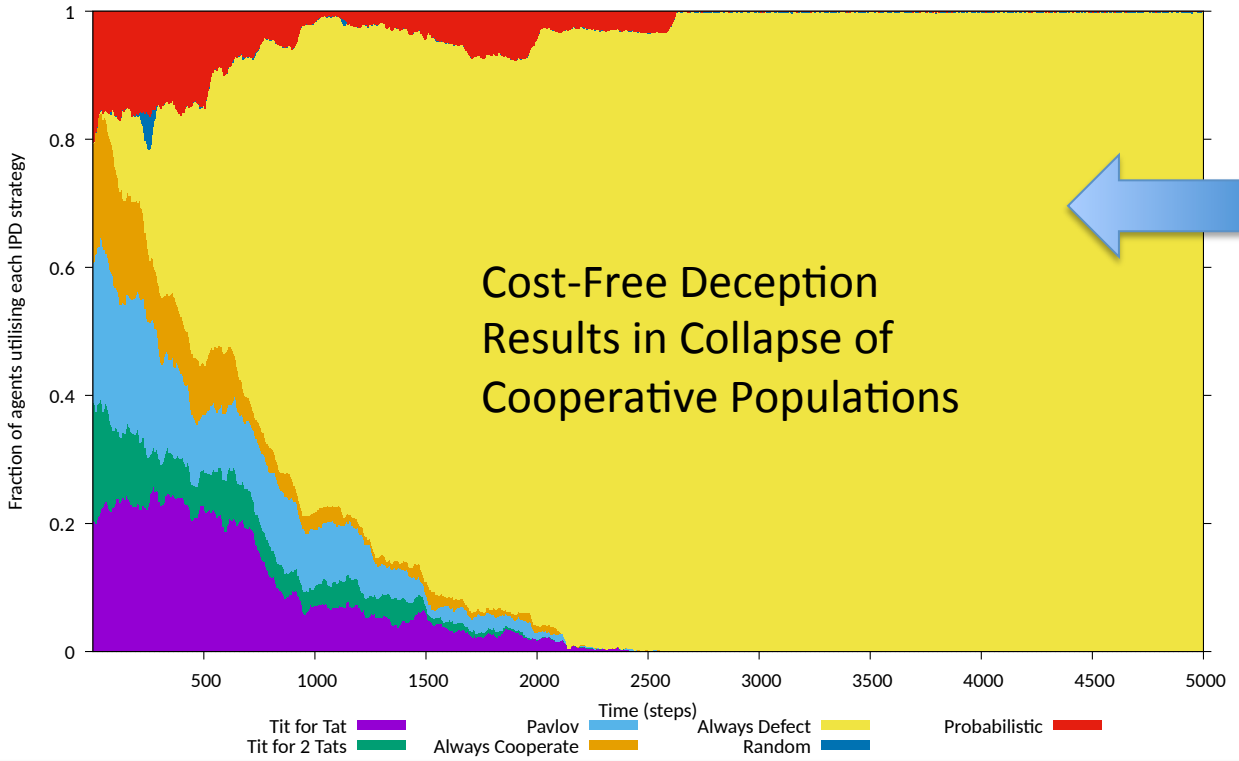


Population of Defecting Agents

Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

Iterated Prisoner's Dilemma Population with Deceiving Players (***Cost 0!***)

Degradation
Deception
Introducing
Uncertainty
And Disrupting
Cooperation

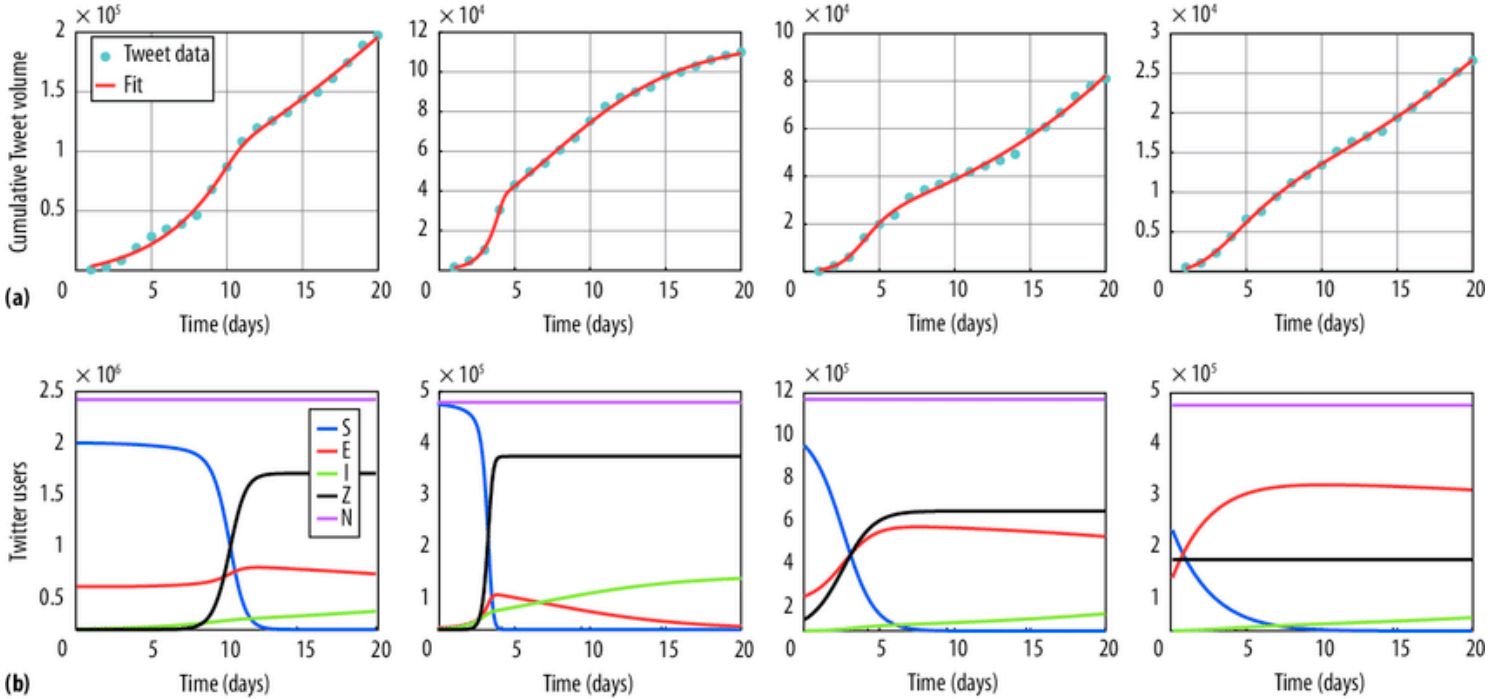


Cost-Free Deception
Results in Collapse of
Cooperative Populations

Population of
Defecting
Agents

Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

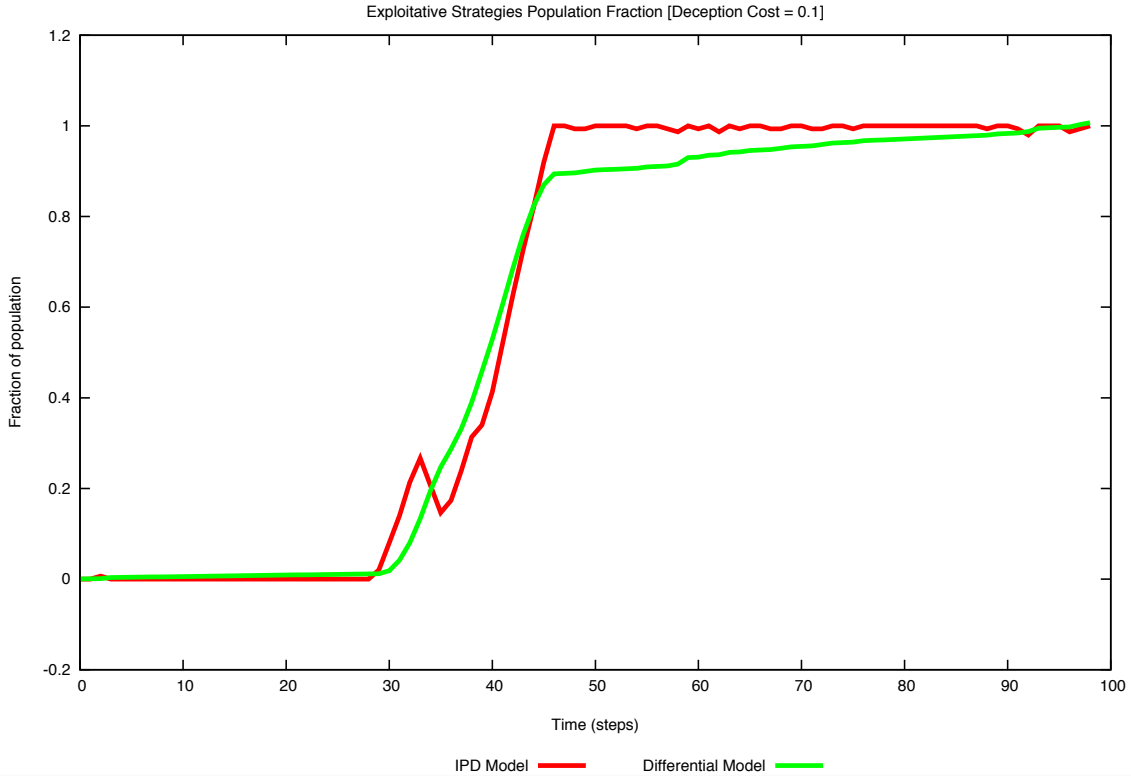
Jin et al 2014 *"Misinformation Propagation in the Age of Twitter"*



Model fits of SEIZ to (a) different rumors and (b) time-course data for each state variable (N denotes the total number of Twitter users). From left to right: White, Zombie, Airborne, and Patent rumors. Jin et al, *"Misinformation Propagation in the Age of Twitter"* , IEEE Computer 47(12):90-94

Diffusion (SIR) in IPD Population with Deceiving Players (Cost 0.3)

Corruption
Deception
Introducing
False Beliefs
And Diffusing
Into Population



Kopp, Korb and Mills, *"Information-theoretic models of deception: modelling cooperation and diffusion in populations exposed to 'fake news'"*, PLOS One, 2018.

Results of the 2016-18 *"Fake News"* Experiments (Monash Uni FIT Study)

- The *Degradation* experiment showed the impact of *Degradation* on consensus forming behaviours in the agent population was very strong – a deceiver population of ~1% could seriously disrupt cooperation;
- The *Corruption* experiment showed the diffusion of *Corruption* in the agent population closely emulated empirical studies of social media;
- Both experiments showed that success, and indeed survival of deceivers in the population strongly depended on the *Cost* of the deception;
- The results of the experiments, despite the simplicity of the simulation, showed good agreement with empirical observations of social media;
- Population behaviours similar to those observed with cheating bacteria in populations of cooperating bacteria (Czaran & Hoekstra, 2009).

Conclusions?

- The *Deception Pandemic* is the result of the confluence of a number of cultural / social trends and Moore's Law driven growth in the digital technology base
- Victims of the *Pandemic* often exist in *"alternate [cognitive] realities"*
- In many ways the *Pandemic* emulates historical patterns seen with Gutenberg's printing press, but many orders of magnitude faster
- Most of the underlying mechanisms are well understood, some less so
- Overcoming the *Deception Pandemic* will not be easy due to the "*information economy*", where deceptive content is too frequently very profitable, frequently wilful behaviour by victims, who often prefer fantasy to fact, and the immense popularity of many of these deceptions in politics

MONASH University

End Presentation

Understanding the Deception Pandemic

GROUP OF EIGHT AUSTRALIA

MONASH University

Backup Slides – Information Theory Primer

Dr Carlo Kopp

Information Theory: Reference Sources and Bibliography

- There is an abundance of websites and publications dealing with basic information theory.
- Examples include:
  - http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html
  - http://okmij.org/ftp/Computation/limits-of-information.html
  - http://www.sveiby.com/articles/information.html
  - http://pespmc1.vub.ac.be/ASC/INFORMATION.html
  - http://www.mtm.ufsc.br/~taneja/book/node5.html
  - http://www.mtm.ufsc.br/~taneja/book/node6.html

Defining Information – Shannon and Weaver

- Shannon – ‘…that which reduces uncertainty…’
- Shannon & Weaver – ‘The quantity which uniquely meets the natural requirements that one sets up for “information” turns out to be exactly that which is known in thermodynamics as entropy.’
- Shannon & Weaver - ‘Information is a measure of one’s freedom of choice in selecting a message. The greater this freedom of choice, the greater the information, the greater is the uncertainty that the message actually selected is some particular one. Greater freedom of choice, greater uncertainty greater information go hand in hand.’ (Sveiby 1994)

Defining Information - Wiener

- Shannon and Wiener define information differently – Shannon sees it as measured by entropy, Wiener by the opposite of entropy.
- Wiener – ‘The notion of the amount of information attaches itself very naturally to a classical notion in statistical mechanics: that of entropy. Just as the amount of information in a system is a measure of its degree of organisation, so the entropy of a system is a measure of its degree of disorganisation.’ (Sveiby 1994)

Wiener Continued

- ‘One of the simplest, most unitary forms of information is the recording of choice between two equally probable simple alternatives, one or the other is bound to happen - a choice, for example, between heads and tails in the tossing of a coin. We shall call a single choice of this sort a decision. If we then ask for the amount of information in the perfectly precise measurement of a quantity known to lie between A and B, which may with uniform *a priori* probability lie anywhere in this range, we shall see that if we put A = 0 and B = 1, and represent the quantity in the binary scale (0 or 1), then the number of choices made and the consequent amount of information is infinite.’

Defining Information - Krippendorf

- Krippendorff - - ‘Literally that which forms within, but more adequately: the equivalent of or the capacity of something to perform organizational work, the difference between two forms of organization or between two states of uncertainty before and after a message has been received, but also the degree to which one variable of a system depends on or is constrained by another. E.g., the DNA carries genetic information inasmuch as it organizes or controls the orderly growth of a living organism. A message carries information inasmuch as it conveys something not already known. The answer to a question carries information to the extent it reduces the questioner's uncertainty.’ …

Krippendorf - Cont

- … ‘A telephone line carries information only when the signals sent correlate with those received. Since information is linked to certain changes, differences or dependencies, it is desirable to refer to theme and distinguish between information stored, information carried, information transmitted, information required, etc. Pure and unqualified information is an unwarranted abstraction. Information theory measures the quantities of all of these kinds of information in terms of bits. The larger the uncertainty removed by a message, the stronger the correlation between the input and output of a communication channel, the more detailed particular instructions are the more information is transmitter.’ (Principia Cybernetica Web).

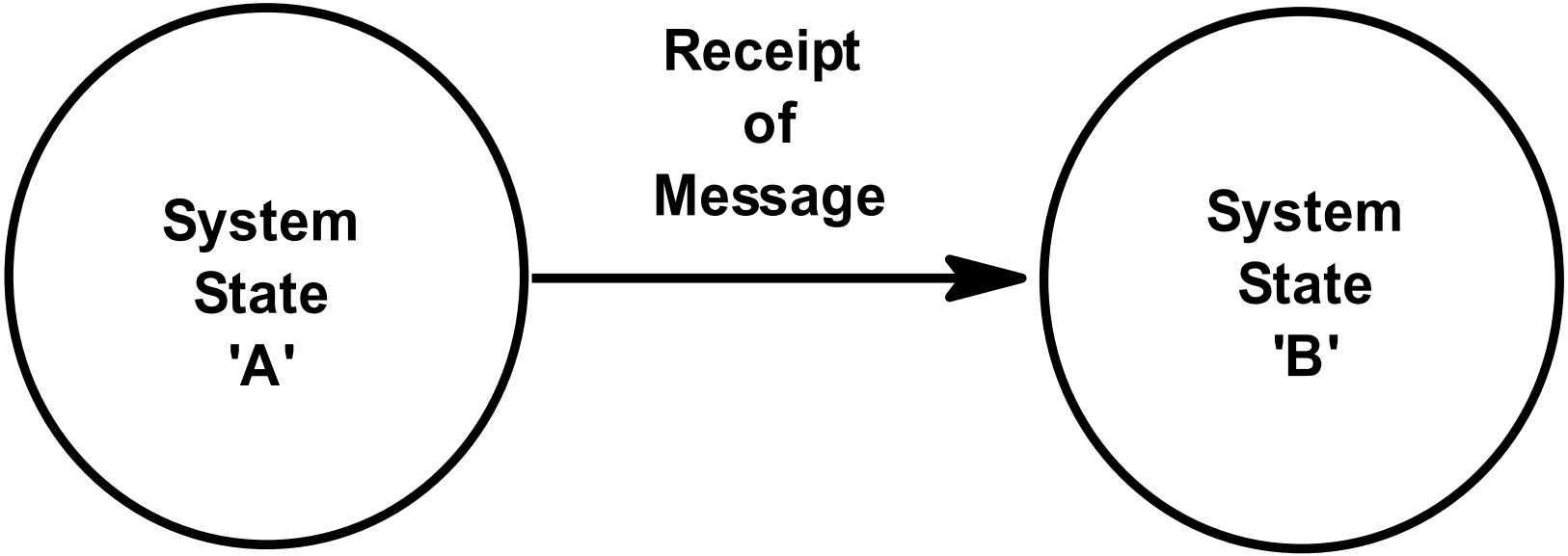
Defining Information - Hornung

- Bernd Hornung - ‘Information is the meaning of the representation of a fact (or of a message) for the receiver.’

Key Issues in Definition

- Information is a means via which the state of uncertainty in an entity can be reduced or changed.
- Entropy in thermodynamics is a measure of the state of disorder in a system; entropy in information theory is a measure of the state of disorder in an information processing system.
- For information to have effect the entity must understand the message in receives; if the message has no meaning to the entity receiving it, it cannot alter the state of uncertainty in that entity.
- If a message which is understood contains information, it will alter the system by changing the state of uncertainty.
- Information can be measured.

State Changes



Meaning in Information

- A key issue which is often not considered in definitions is the matter of *meaning* – can the message be understood?
- A work of Shakespeare written in English will be rich in information content, but only to an English speaker.
- An English speaker with good knowledge of Elizabethan English will perceive greater information content than a reader without; a reader with better knowledge of period history will perceive greater information content than a reader without; and so on …
- *In mathematical terms, the receiver of the message must be capable of decoding the message, to determine what information it contains.*

Example

- DNA encodes the definition of an organism's structure and function.
- Alter the DNA chain of an embryo and the resulting organism will be different, possibly in many ways.
- Does this mean that we can splice DNA in any manner we choose?
- For the DNA to be properly decoded, it needs to be inside a biological entity which can process (understand) what the DNA tells it to do.
- If the species between which the DNA is being spliced are too different, the DNA is not likely to be decoded in the manner intended, resulting in a non-viable organism.

Shannon's Entropy

- Shannon defines entropy as a measure of uncertainty, where *H* is entropy, and p_i is the probability of a symbol or message (Theorem 2):

$$H(X) = -\sum_{i=0}^{N-1} p_i \log_2 p_i$$

- The logarithm is base 2.
- Shannon's proof is well worth reading – refer:

  'Properties of Shannon's Entropy'
- This is based on the paper 'A Mathematical Theory of Communication'

Thermodynamics - Entropy

- The second law of thermodynamics ie '*the total entropy of any isolated thermodynamic system tends to increase over time, approaching a maximum value*' is often represented as:

$$S = -k_B \sum_i p_i \log_2 p_i$$

- Where k_B is Boltzmann's constant or k = 1.3806505 x 10^{-23} [joules/kelvin].
- This form is for all intents and purposes the same as that proven by Shannon in the Entropy Theorem.

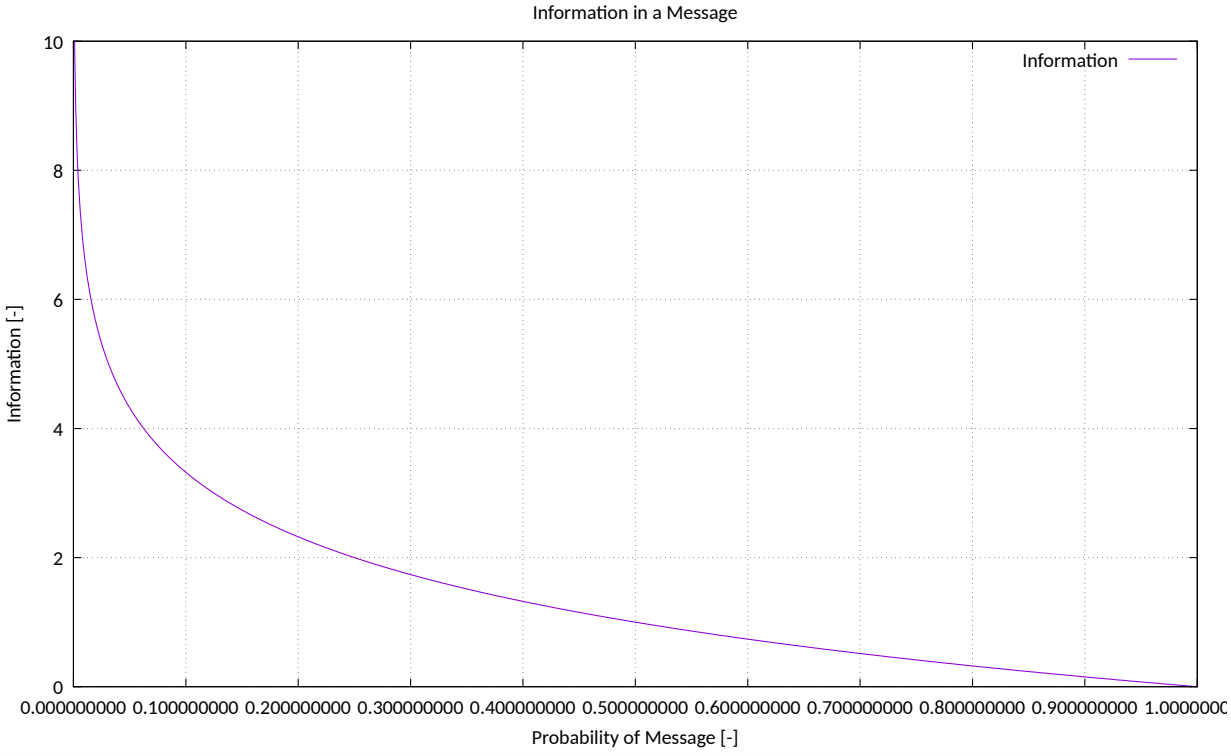
Information in a Message

- The Entropy Theorem presents the total entropy (information) in the system, across all of the possible messages in the system.
- For many problems we are interested in the information in one of the N messages. This can be represented thus:

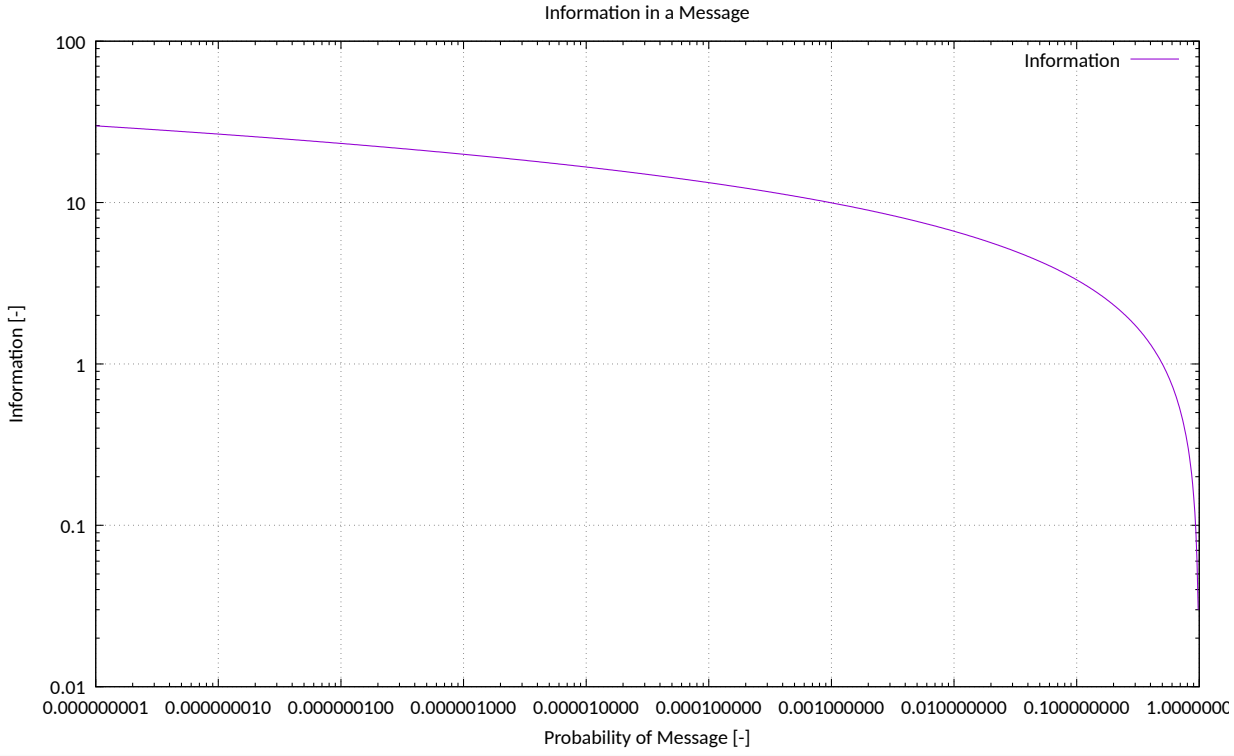
$$I(m) = -\log_2(p(m))$$

- As is evident, highly probable messages contain little information and vice versa.

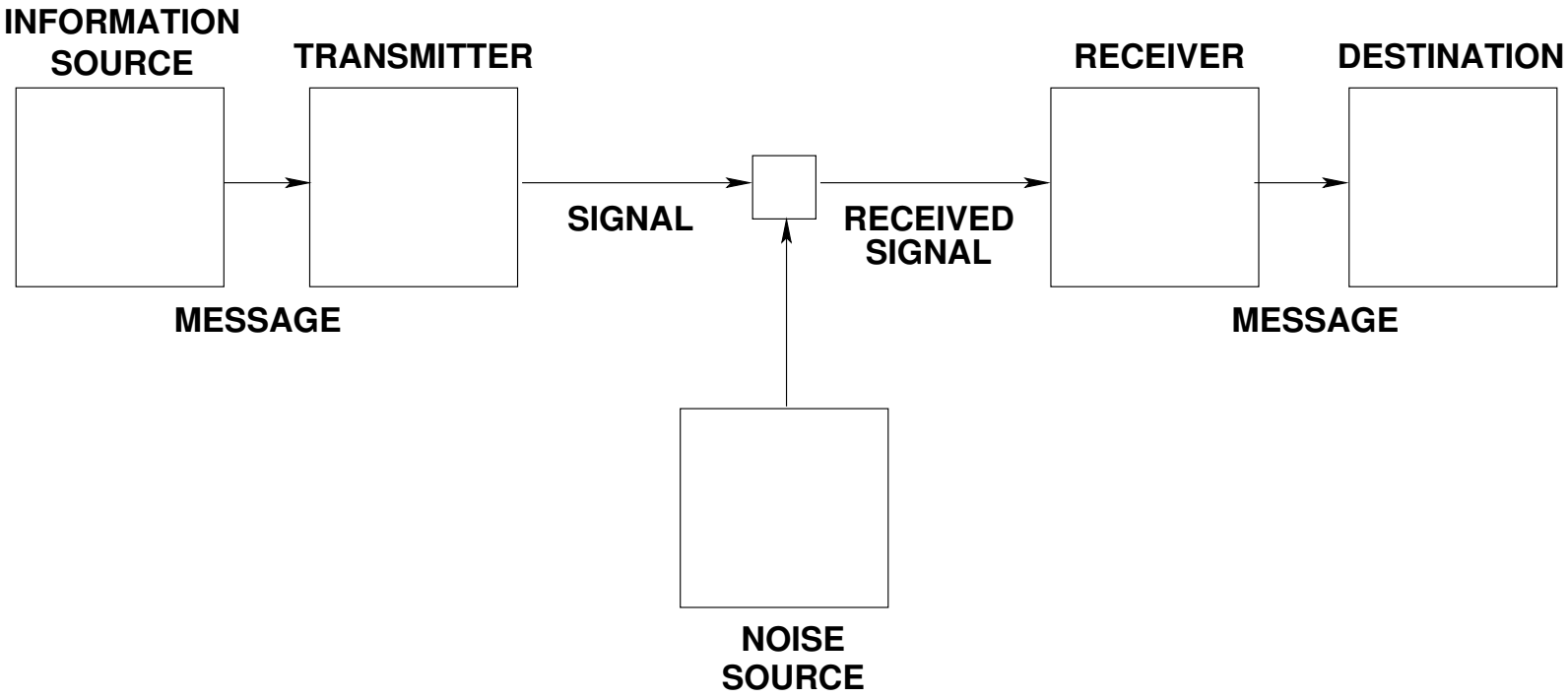
Information vs Likelihood of Message



Information vs Likelihood of Message



Shannon's Channel Capacity Model (1)



Shannon's Channel Capacity Model (2)

The model has five key components:

1. The 'information source' which generates messages containing information
2. The 'transmitter' which sends messages over the 'channel'.
3. The 'channel' and associated 'noise source', this could be any number of physical channel types including copper or optical cable, radio link or acoustic channel.
4. The 'receiver' which detects and demodulates messages received over the 'channel'.
5. The 'destination' or 'information sink' which responds to messages by changing its internal state.

It is implicitly assumed that messages sent by the 'information source' can be understood by the 'sink.

Shannon's Channel Capacity Model (3)

- Shannon demonstrated that for a 'noisy' channel, ie one in which random noise could additively contaminate the messages flowing across the channel, the capacity of the channel (amount of information it could carry) is defined by (Theorem 17):

$$C = W \log_2 (\frac{(P+N)}{N})$$

- Where *C* is channel capacity, *W* is channel bandwidth, *P* is signal power, and *N* is noise power.
- This equation is most commonly used in the following form, as *P/N* is the widely used measure of 'signal to noise ratio' or 'SNR':

$$C = W \log_2 (1 + \frac{P}{N})$$

Shannon's Channel Capacity Model (4)

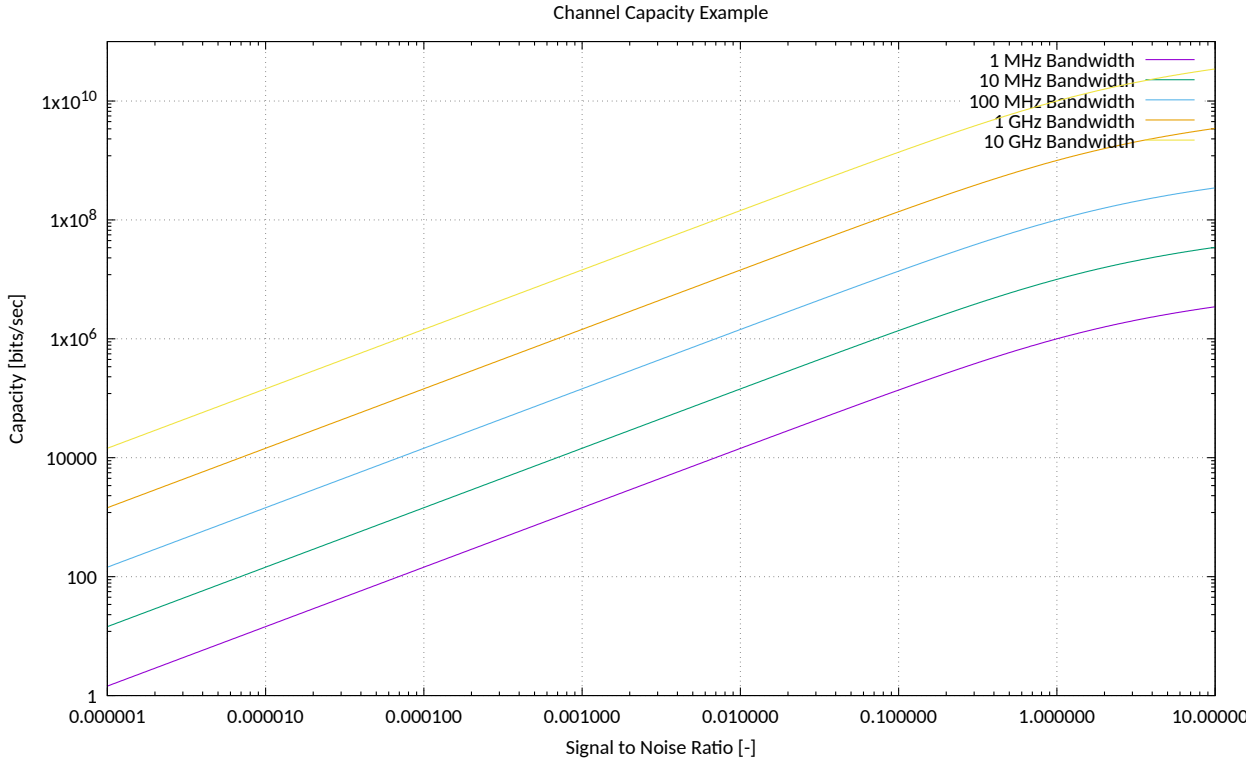
- Assumption (1) – the additive noise is 'white thermal noise' ie it has a normal distribution in time/power;
- Assumption (2) – the power of the signal (message) in the channel is the average, rather than peak power;
- Assumption (3) – the power is not limited by the transmitter's peak power rating;
- Metrics (1) – channel capacity is defined in bits/sec;
- Metrics (2) – bandwidth is defined in Hertz (cycles/sec);
- Metrics (3) – signal and noise power are defined in Watts;
- In numerical applications which compute capacity, it is customary to use this form as $\log_2$ is often not available:

```
C=B*(1/log(2))*log( 1 + S/N );
```

What Does This Model Tell Us?

- It is possible to trade between bandwidth and signal to noise ratio to achieve an intended capacity – an example is in spread spectrum communications;
- Channels with severely limited bandwidth but very high signal to noise ratio can still achieve high capacity – example are voiceband modems running over digital switches;
- Where the SNR >> 1, the second term approximates the logarithm of SNR;
- Where the SNR << 1, the second term -> 0, and bandwidth becomes the dominant means of improving capacity;
- *By manipulating the bandwidth and SNR of a channel we can manipulate its capacity, and thus how much information it can carry.*

Channel Capacity Example



Key Points

- What constitutes information in a message depends on the ability of an entity to understand that information.
- If a message contains information, an entity receiving it and understanding it will experience a state change which alters its level of uncertainty.
- The less likely the message, the greater its information content (Entropy theorem).
- The capacity of a channel to carry information depends on the magnitude of interfering noise in the channel, and the bandwidth of the channel (Capacity theorem).

MONASH University

Backup Slides – Deception Model Primer

Dr Carlo Kopp

Reference Sources and Bibliography

- There are only two primary sources dealing with the four canonical strategies:

1. Borden, Andrew; What Is Information Warfare? *Air & Space Power Chronicles*, November 1999.
2. Kopp, Carlo; A Fundamental Paradigm of Infowar, *Systems*, February, 2000.

- Supporting definitions can be found in: United States Dept of the Air Force; Cornerstones of Information Warfare; Washington, 1995. 13 p. also at http://www.c4i.org/cornerstones.html

Background to the Four Canonical Models

- The four canonical models of deception were identified almost concurrently by Colonel Andrew Borden, PhD, USAF, and Carlo Kopp, at Monash University CSSE, in 1999.
- Dr Borden published two months before Kopp in APJ *Air Chronicles*, a United States Air Force professional journal. Kopp published in the Australian industry journal *Systems*, formerly *Australian Unix User's Review*.
- Borden's model does not include the '*subversion*' strategy as a defined model, and follows the US DoD convention of opaquely conflating it into the '*denial*' model.
- The *subversion* model was first published by Kopp, and credit for its identification must go to the late Prof C.S. Wallace, foundation Chair of Computer Science at Monash University.

Why a Fundamental Theory/Paradigm?

- Prior to the definition of the Borden-Kopp model for Information Warfare, there was no established mathematical basis to underpin the theory.
- As a result considerable disagreement emerged in the literature and professional debate as to even the basic validity of the idea of information use in survival conflicts.
- With the definition of a mathematically supportable and robust theoretical basis, this area of study can now be explored scientifically and in a systematic fashion.
- Subsequent research has described the relationship between games and information, and the properties of compound strategies.
- Later research by Kopp and Mills also established the role of information conflicts in biological evolution.

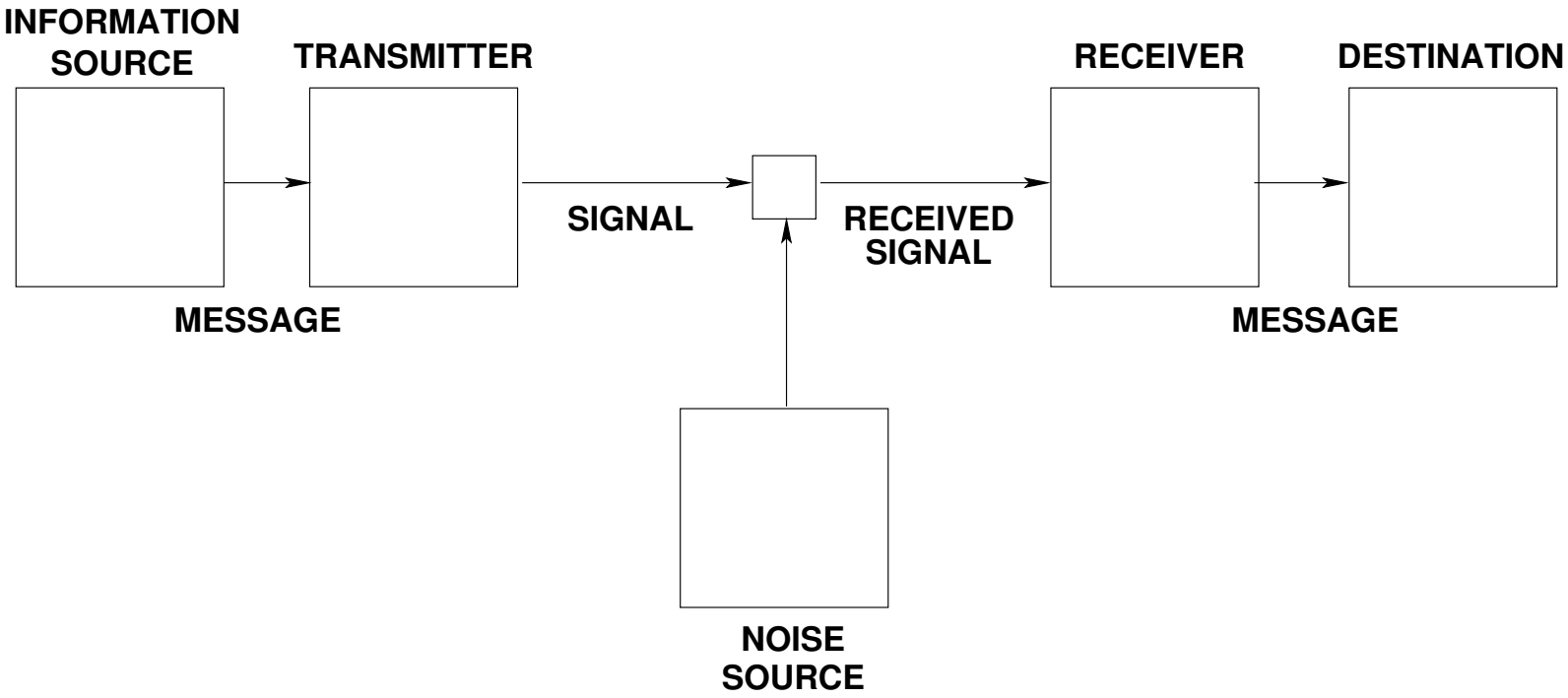
The Starting Point - Shannon's Capacity Model

- To establish a fundamental theoretical model the starting point must be fundamental information theory, which is centred in Shannon's channel capacity theorem:

$$C = W \log_2(1 + \frac{P}{N}) \quad \text{Theorem 17}$$

- If an attacker intends to manipulate the flow of information to an advantage, the game will revolve around controlling the capacity of the channel, *C*.
- To achieve this, the attacker must manipulate the remaining variables in the equation, bandwidth, *W,* and signal power vs noise power, *P/N.*
- *Three of the four canonical models involve direct manipulation of bandwidth, signal power and noise.*

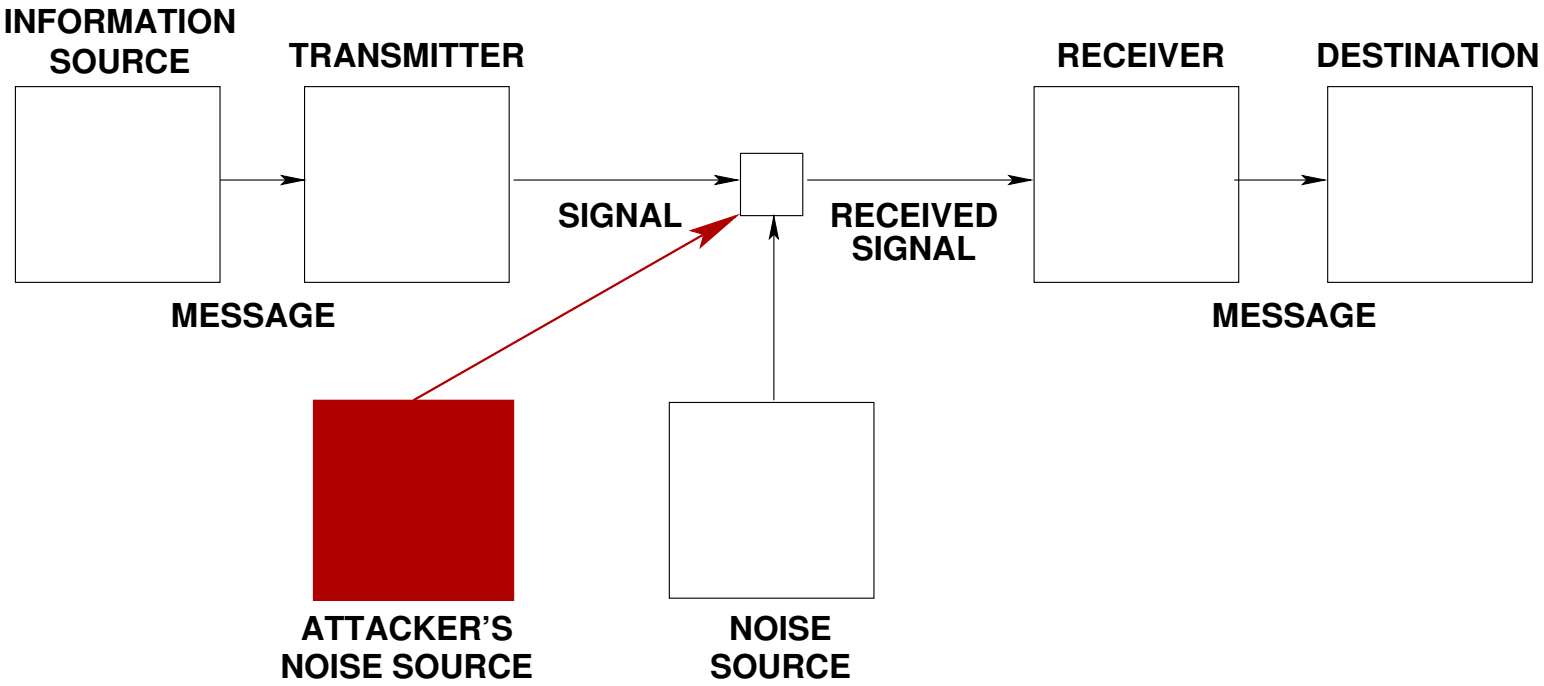
Shannon's Channel Capacity Model



The Degradation Deception

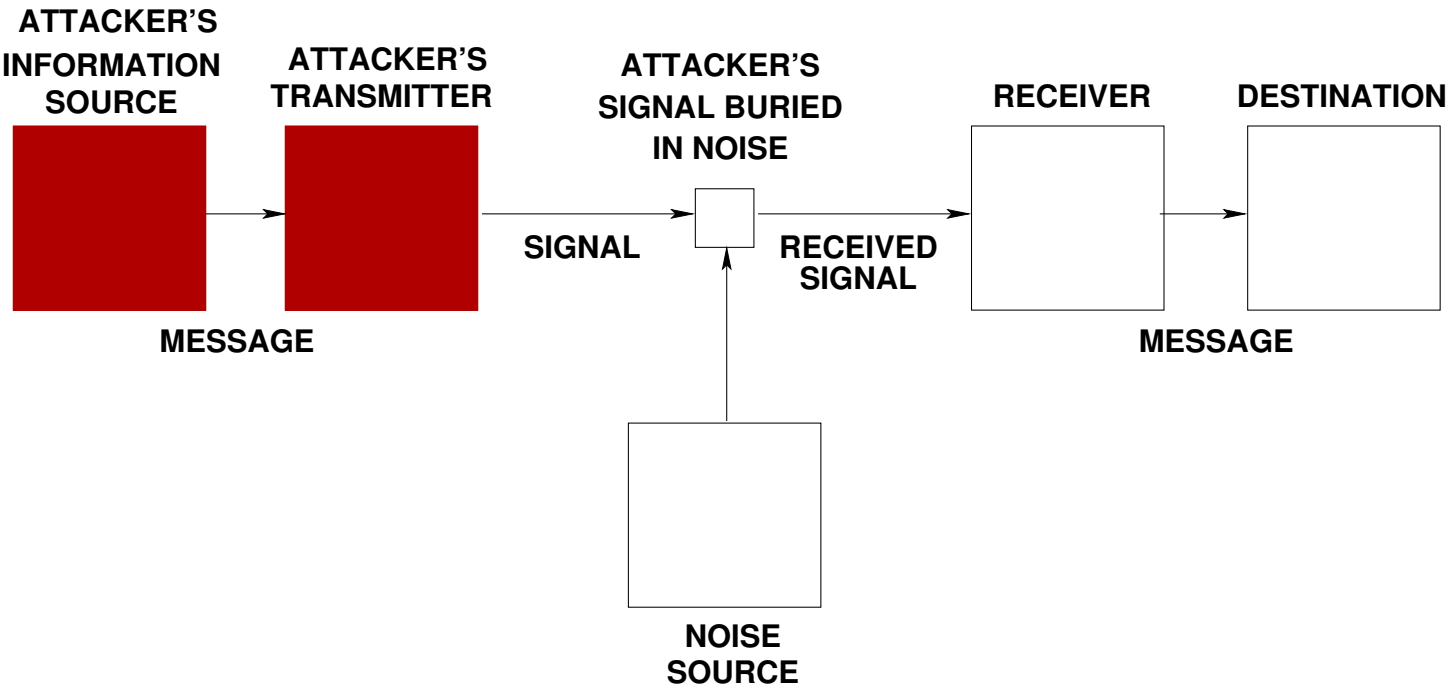
- The degradation deception involves manipulation of the P/N term in Shannon's equation.
- The flow of information between the source and destination is impaired or even stopped by burying the signal in noise and driving $C \rightarrow 0$.
- There are two forms of this strategy, the first being the 'covert' or 'passive' form, the second being the 'overt' or 'active' form.
- The first form involves forcing $P \rightarrow 0$ to force $C \rightarrow 0$. In effect the signal is made so faint it cannot be distinguished from the noise floor of the receiver.
- The second form involves the injection of an interfering signal into the channel, to make N>>P and thus force $C \rightarrow 0$. In effect the interfering signal drowns out the real signal flowing across the channel.

Degradation Model – Overt Form



Degradation Deception Model

Degradation Model – Covert Form



Degradation Covert / Passive Deception Model

Passive / Covert vs Active / Overt Forms of Degradation

- There is an important distinction between the active / overt and passive / covert forms of the degradation deception.
- In the *passive form* of this attack, the victim will most likely be unaware of the attack, since the signal is submerged in noise and cannot be detected. This form is therefore '*covert*' in the sense that no information is conveyed to the victim.
- In the *active form* of this attack, the signal which jams or interferes with the messages carried by the channel will be detected by the victim. Therefore this form is '*overt*' in the sense that information is conveyed to the victim, telling the victim that an attack on the channel is taking place.
- Both forms are widely used in biological survival contests and in social conflicts.

Example - Degradation

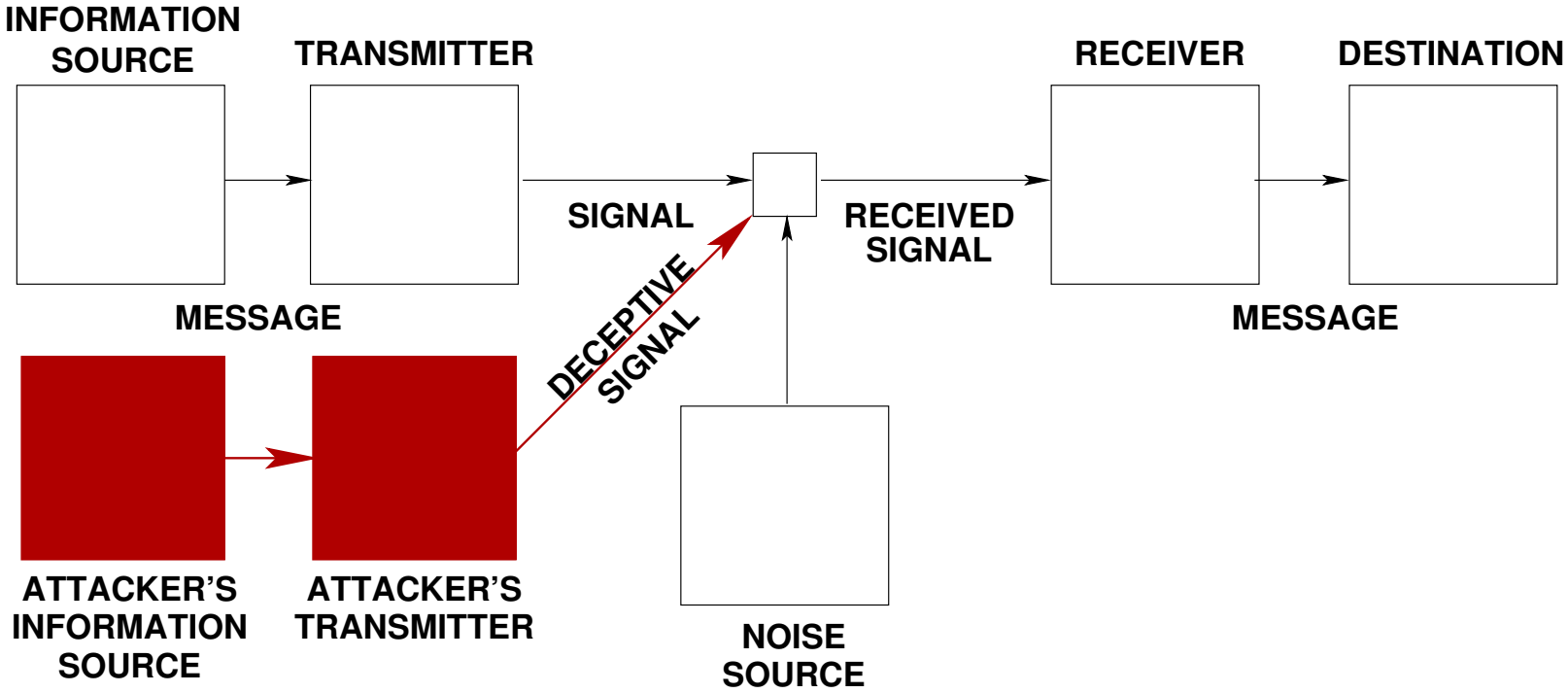
Examples - Degradation

- Passive form – biological or military camouflage patterns.
- Passive form – military stealth to hide from radar.
- Passive form - encryption and concealment to prevent unwanted parties from reading or finding what they ought not to.
- Active form – barrage jamming of wireless radio broadcasts or communications links.
- Active form – the use of smoke screens to hide troops from enemy gunfire.
- Active form – biological examples such as squid squirting ink at predators to hide themselves.

Corruption Deception Model [Mimicry]

- The corruption model involves the substitution of a valid message in the channel with a deceptive message, created to mimic the appearance of a real message.
- In terms of the Shannon equation, S_{actual} is replaced with S_{mimic}, while the *W* and *N* terms remain unimpaired.
- The victim receiver cannot then distinguish the deception from a real message, and accepts corrupted information as the intended information. Success requires that the deceptive message emulates the real message well enough to deceive the victim.
- Corruption is inherently 'covert' since it fails in the event of detection by the victim receiver. Corruption is used almost as frequently as degradation in both biological and social conflicts.

Corruption Deception Model



Corruption Deception Model

Example - Corruption

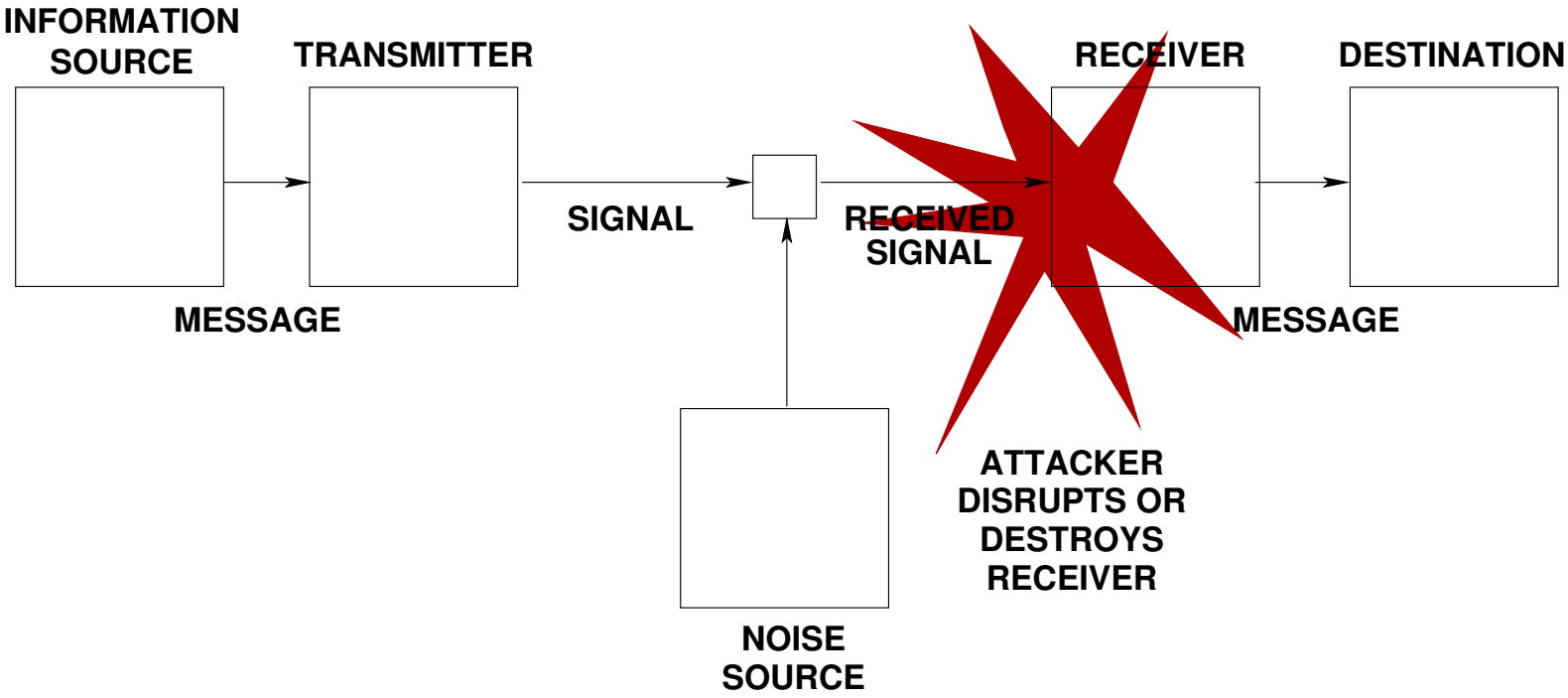
Examples - Corruption

- Biological examples of organisms which mimic the appearance of harmful, predatory or toxic species to deceive predators.
- Biological predators which mimic the appearance of prey organisms to attract lesser predators and eat them.
- Deception jamming techniques used against radars, producing errors in angle/range measurements, or producing false (non-existent) targets.
- The use of deceptive propaganda radio broadcasts, or deceptive radio transmissions emulating real messages.
- Deceptive advertising in the commercial and political domains.
- Identity theft, phishing, phracking, hacker use of stolen usercodes, spammer email address substitution.

Denial Deception Model

- The degradation and corruption strategies both focus on the *P* and *N* terms in the Shannon equation.
- The denial strategy manipulates the *W* term, by effecting an attack on the transmission link or receiver to *deny* the reception of any messages, by removing the means of providing bandwidth *W*.
- This means that $W \rightarrow 0$ or $W = 0$ if the attack is effective.
- The denial strategy is inherently 'overt' in that the victim will know of the attack very quickly, as the channel or receiver is being attacked.
- A denial attack may be temporary or persistent in effect, depending on how the channel or receiver is attacked.
- Numerous biological and social examples exist.

Denial Deception Model



Example – Denial

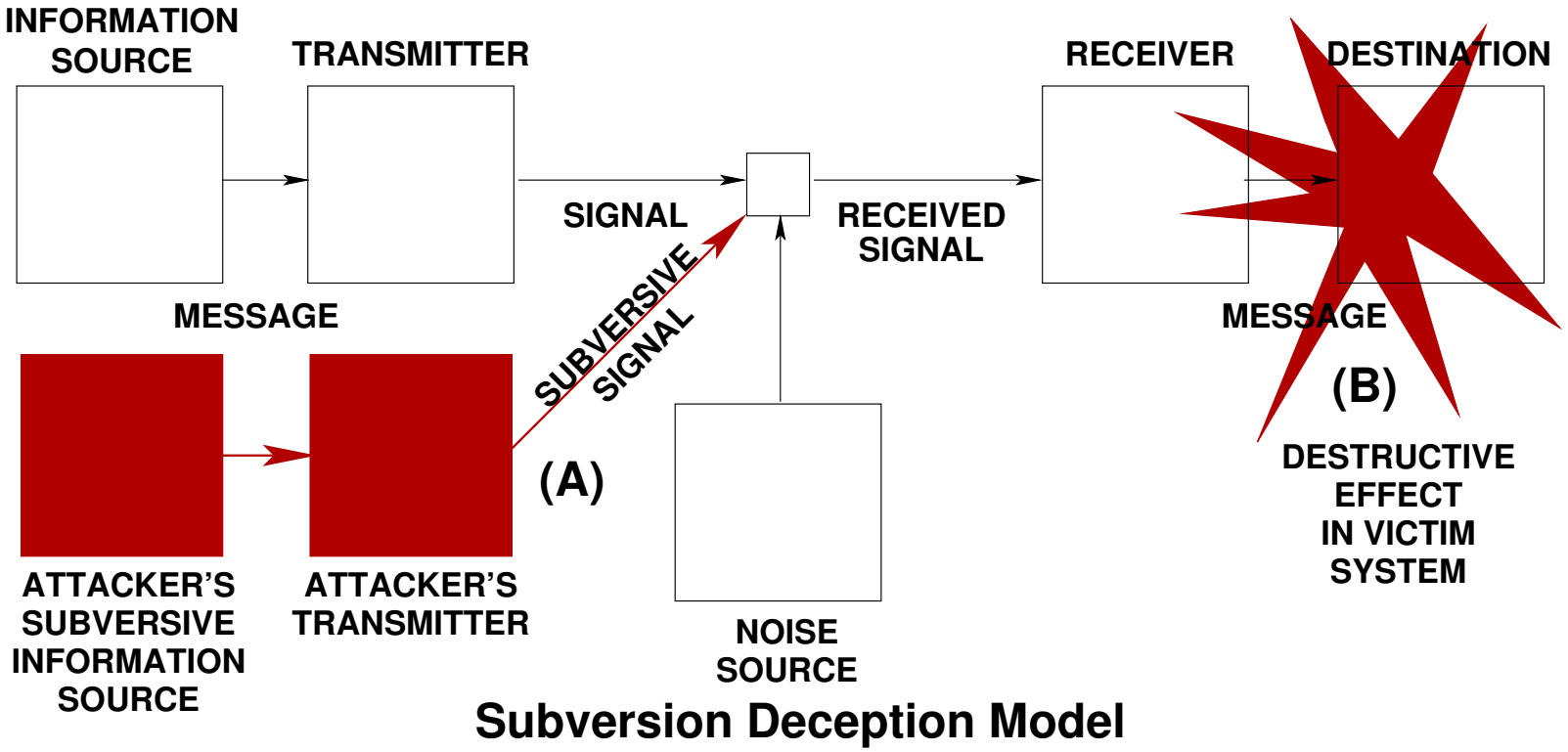
Examples – Denial

- Organisms which spray noxious fluids on predators, thereby blinding and numbing the predator's visual and olfactory senses, temporarily or permanently.
- Very high power radio frequency weapons which can permanently or temporarily impair the function of victim receivers by overloading input circuits.
- Destroying the receiver system by direct attack, for instance by fire, bombing or other such means.
- In the IT domain, any temporary or permanent ‘denial of service’ attack, such as ‘ping of death’, induced packet storms, cutting data or power cables, or using electromagnetic weapons.

Subversion Deception Model

- Subversion differs from the first three models in that it does not involve an attack on the message, its contents or the channel/receiver.
- Subversive attacks involve the insertion of information which triggers a self destructive process in the victim system or organism.
- At the most basic level this is the diversion of the thread of execution within a Turing machine, which maps on to the functional behaviour of the victim system.
- It amounts to surreptitiously flipping one or more specific bits on the tape, to alter the behaviour of the machine.
- The attack may impair or destroy the victim system.
- Numerous biological, social and technological examples exist.

Subversion Deception Model

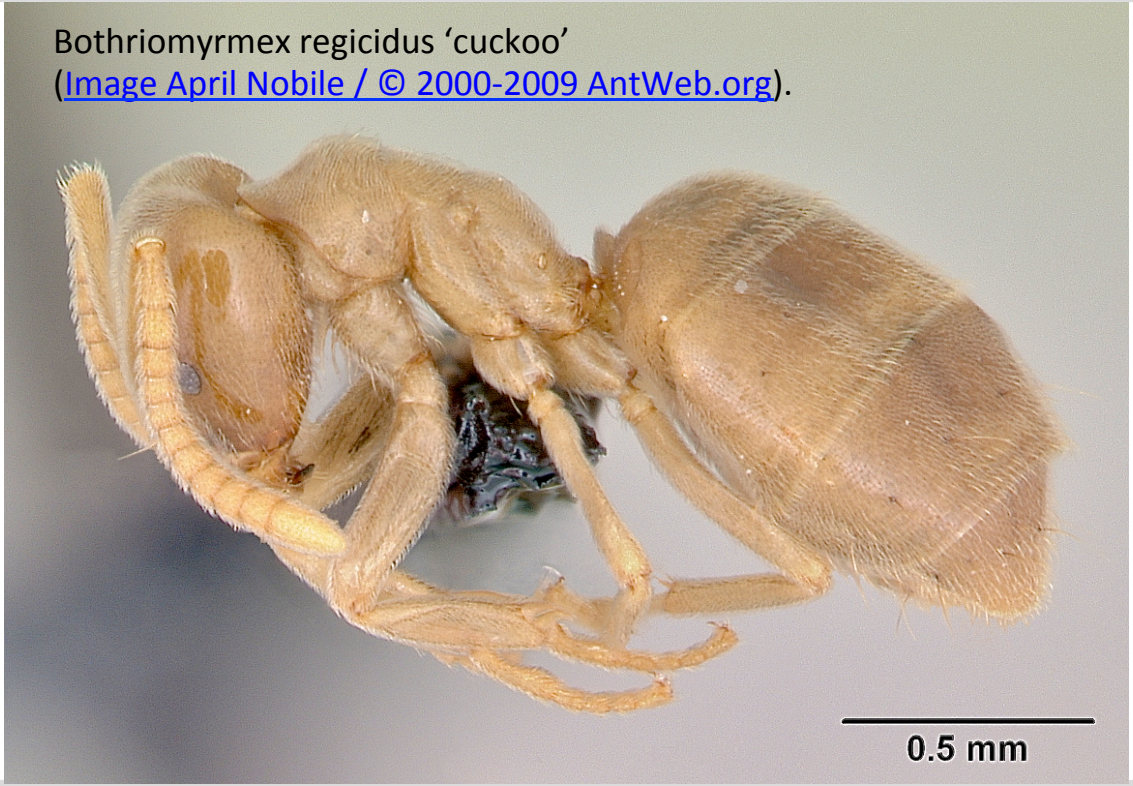


Examples - Subversion

- Parasites which emit chemicals which alter the internal functions of the victim organism to favour the parasite, such as the production of favourable nutrients or weakening of immune defences.
- The use of deceptive radio or optical signals which trigger the premature initiation of weapon fuses, such as proximity fuses on guided missiles or artillery shells.
- Logic bombs, viruses, worms and other destructive programs which use system resources to damage the system itself.
- Most examples of subversion rely on the attacker's use of corruption to penetrate the victim's defences and create conditions to effect the subversive attack.

Examples - Subversion

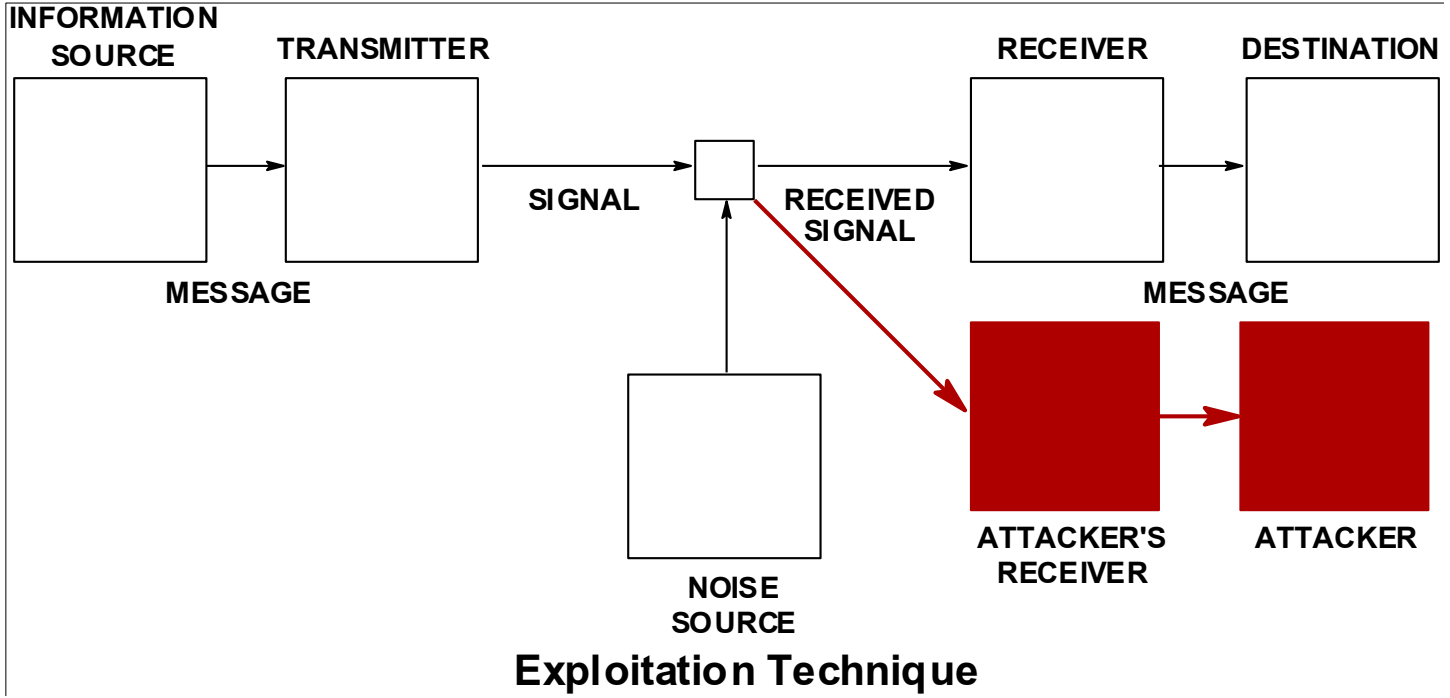
Bothriomyrmex regicidus ‘cuckoo’
(Image April Nobile / © 2000-2009 AntWeb.org).



Why Exploitation is Not a Canonical Model

- The US DoD definitions of the four strategies of information attack include ‘exploitation’, which is effectively the eavesdropping of victim messages.
- As eavesdropping is a wholly passive activity which does not involve a direct attack on the victim channel, receiver or system, thus impairing or altering the function of the victim, it cannot be a canonical strategy defining a mode or type of attack on a system.
- For completeness, exploitation is defined and illustrated.

Exploitation



Proving the Four Canonical Strategies (Models)

- Early critics argued that IW did not exist and had no scientifically provable basis (none of them were scientists!).
- ***Proof****: If IW does not exist as an artifact of evolution in nature, then no examples of its use should exist. As examples exist in abundance, then this hypothesis is clearly false.*
- Do other possible canonical models exist?
- There are only three variables in the Shannon equation, each accounting for one of the first three strategies. In a Turing machine, information can be used to alter the program but not the nature of the machine.
- *Hence, there are no obvious candidates for further canonical models*.

Properties of the Four Canonical Models

- **Orthogonality**: A canonical model cannot be formed by combining any number of the remaining canonical models. **Proof**: each strategy attacks the victim system in different ways.
- **Indivisibility:** Canonical models cannot be further divided or decomposed. **Proof**: Each of the canonical models represents the simplest way to effect their respective modes of attack.
- **Concurrency:** A victim system can be subjected to any number of concurrent attacks. **Proof**: For *like* attacks, the effects on the victim system are additive; for dissimilar attacks, the effects on the victim system are *orthogonal*.

Nomenclature

US Department of Defense Nomenclature (1995)	**Monash University Nomenclature (1999)**
Degradation	Denial of Information (DoI)
Corruption	Deception and Mimicry (D&M)
Denial	Disruption & Destruction (D&D)
Denial	Subversion (SUB)
Exploitation	N/A

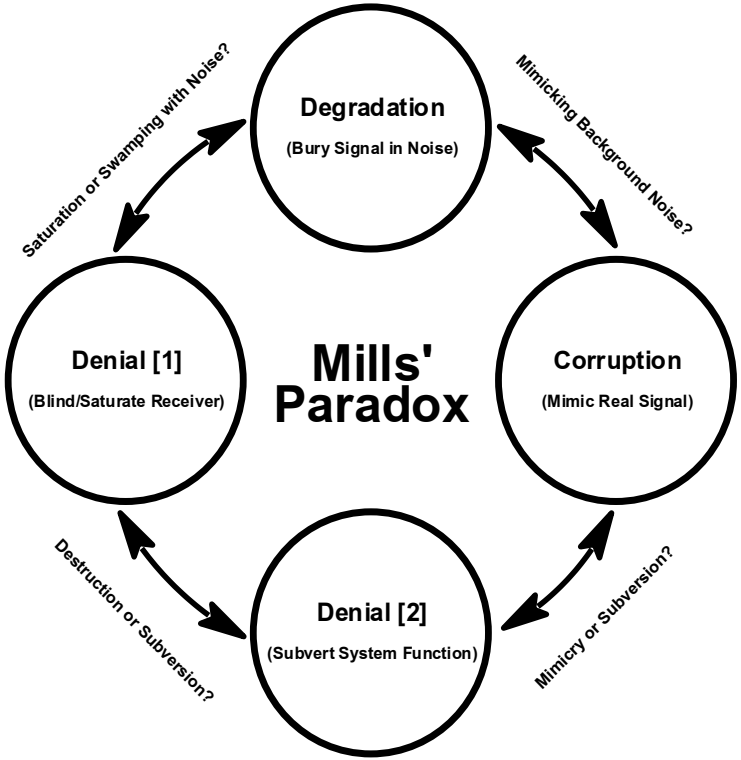
Key Points

- The four canonical models define all modes of attack involving information in terms of basic manipulation of fundamental models – the Shannon channel model and the Turing machine.
- All attacks on information processing or transmission systems comprise either a canonical model or some combination of canonical models.
- The canonical models are ubiquitous in the biological and social domains.
- The four canonical models provide a mathematically robust and provable model for conflicts involving the use of information.

Mills' Paradox

- First identified in 2002 by Mills.
- How do we distinguish a *Denial* via subversion attack from a *Corruption* attack? How do we distinguish a destructive *Denial* via subversion attack from a *Denial* via destruction attack?
- How do we distinguish a *Degradation* attack from a mimicking *Corruption* attack? How do we distinguish an intensive active *Degradation* attack from a soft kill *Denial* via destruction attack?
- *Note that Degradation attacks can always be easily distinguished from Denial via subversion attacks, and Corruption attacks can easily be distinguished from Denial via destruction attacks.*

Mills' Paradox



MONASH University

Backup Slides – Compound Attacks

Dr Carlo Kopp

Reference Sources and Bibliography

- There is only a single reference covering compound information conflict strategies:
- Kopp, Carlo, The Analysis of Compound Information Warfare Strategies, Conference Paper, *Proceedings of the 6th Australian Information Warfare & Security Conference 2005*.

Compound Information Attacks?

- *A compound information attack is any attack that comprises more than one canonical deception model, and in which some defined precedence relationships exist between these strategies.*
- Such attacks arise very frequently in biological and social contexts.
- Empirical study of examples indicates that such attacks can have very large numbers of components.
- The analysis of any such attacks can present difficulties in the absence of systematic techniques for analysis.
- The *de facto* orthogonality property of the canonical strategies, and the existence of precedence relationships permit systematic analysis.

Problems?

- Understanding and analysing a complex compound deception attacks. Such an attack can comprise a very larger number of canonical primitives.
- Properly understanding the structure of the attack, and thus its underlying aims, can present difficulties.
- Example: an opponent is playing a very complex compound deception attack. The aim of the defender is to determine whether gathered information is a deception or not, and what the specific aim of that deception might be. In the simplest of terms, 'what does this opponent want me to think and why?'
- Detection of inconsistencies, mistakes or gaps in such a complex deception strategy may be the only method of unmasking such a deception, especially if the deception is carefully architected from the outset.

Problems? (Continued)

- Another problem which can frequently arise is that of countering an opponent's deceptive perception management attack.
- Such deceptions can often be complex compound attacks in which multiple mutually reinforcing falsehoods are employed with a specific aim of shifting the perceptions of a victim audience.
- Often the only technique for defeating such a strategy is to unmask the deception before the audience.
- A well crafted compound attack may present genuine difficulties in analysis and defeat.

Primitives , Precedence, Compound Strategies

- **The Attacker:** the player in an information warfare attack who is executing the strategy against a victim player.
- **The Victim:** the player in an information conflict attack who is being subjected to an attack by the attacking player.
- **Canonical Model:** defined as one of the four fundamental models. These models are atomic, in the sense that any compound model can be divided into a number of canonical models, but a canonical model cannot be further divided in any way.
- **Compound Strategy:** any attack which comprises more than one canonical information conflict model, and in which some defined precedence relationships exist between these models.

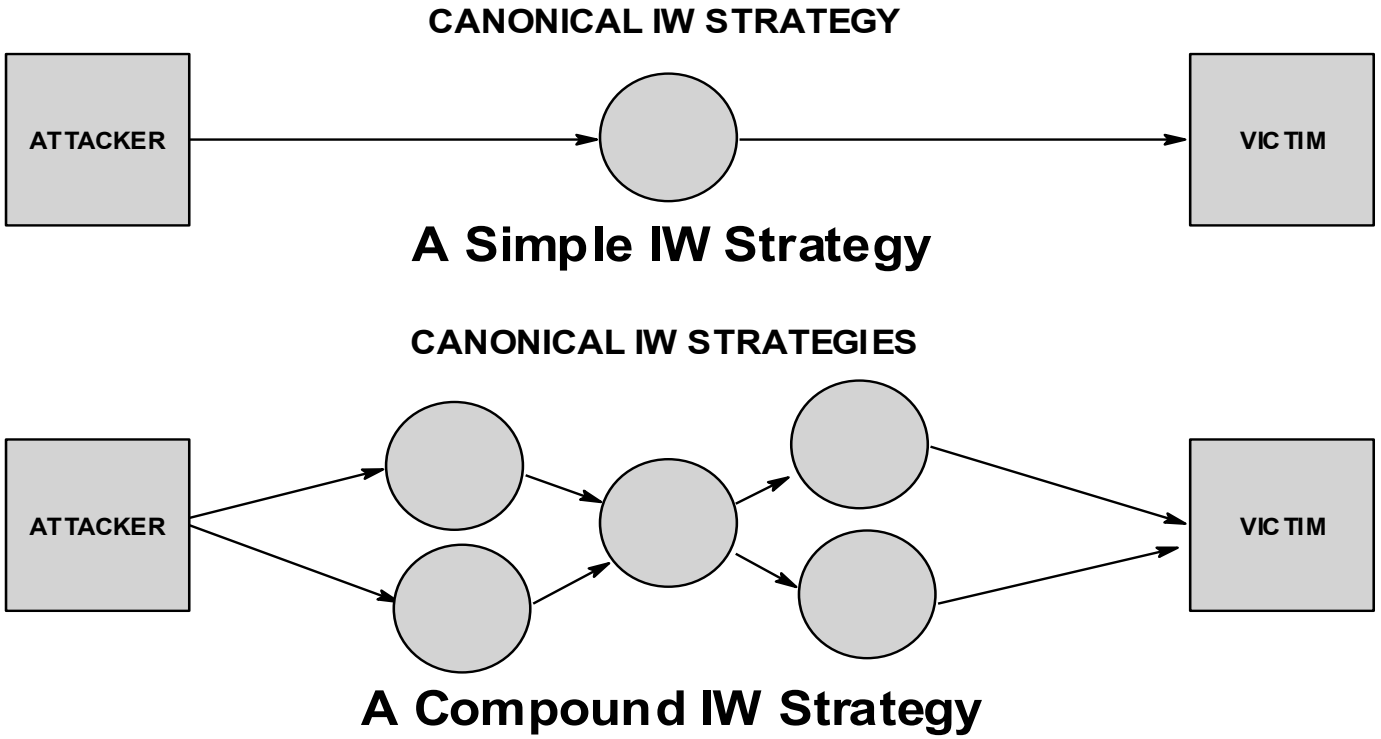
Precedence Relationships

- **Precedence Relationships:** define the order or precedence which exists between more than one canonical information conflict model comprising a compound model:
  1. **In practical terms, one canonical model can be a precedent to one or more canonical models.**
  2. **The precedence relationship cannot be bidirectional since the time domain is not bidirectional (i.e. digraphs).**
  3. **It is only once the precedent strategy has achieved some effect, that the antecedent model can produce its effect.**
  4. **There is no bound on the number of precedent model to any antecedent model.**

Precedence Relationships (Cont)

5. **Precedence is unidirectional in time, therefore any compound model forms a directed graph, which obeys the properties of directed graphs.**
6. **Precedence relationships arise due to the state of the victim in the attack. In a compound model, antecedent strategies may not be feasible until a specific state of misperception or false belief has been established in the victim. A model may only be successful if this state change has taken place.**
7. **An attacker may or may not perceive the state change in the victim's perception arising from an attack, compound or simple, and thus execute an antecedent model, compound or simple, after executing the precedent attack. This may or may not impair the success of the antecedent attack.**

Simple vs Compound IW Strategies



Primitives (Cont)

- **Concurrency:** Attacks between which no precedence relationship exists can be executed concurrently. There is no bound on the number of possible concurrent attacks.
- **Primary vs Supporting Strategies:** An attack is said to be a supporting attack if it supports the aim of another attack, termed the primary attack.
-

1. **Supporting and primary attack may or may not be concurrent.**
2. **A non-concurrent supporting attack is a attack which must produce its effect before the primary attack can be executed successfully.**

Primitives (Cont)

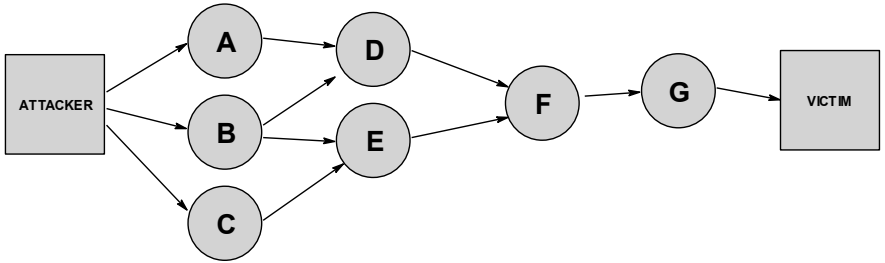
- **Chained or Sequential Attacks:** a compound attack in which one or more intermediate victims are exploited. In such an attack the first victim is employed as a conduit or proxy to propagate an information conflict attack, or its effect.

  - **Example: exploitation of media organizations by terrorist movements. The media organization is deceived into propagating a message targeted at a victim population, believing the message constitutes legitimate news.**

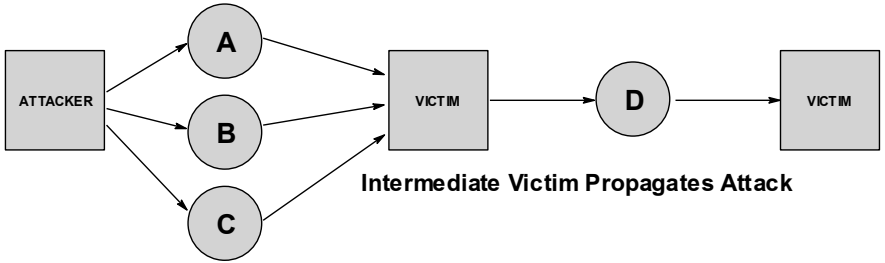
Primitives (Cont)

- **Victim State:** defined as the victim's belief at that point in time.
  - **A successful application of information attack will effect an intended state change.**
  - **An unsuccessful application may not produce a state change, or may by alerting the victim, produce a state change in whatever other game the victim may be playing**.

Chained Compound vs Compound Attacks



A Compound IW Strategy

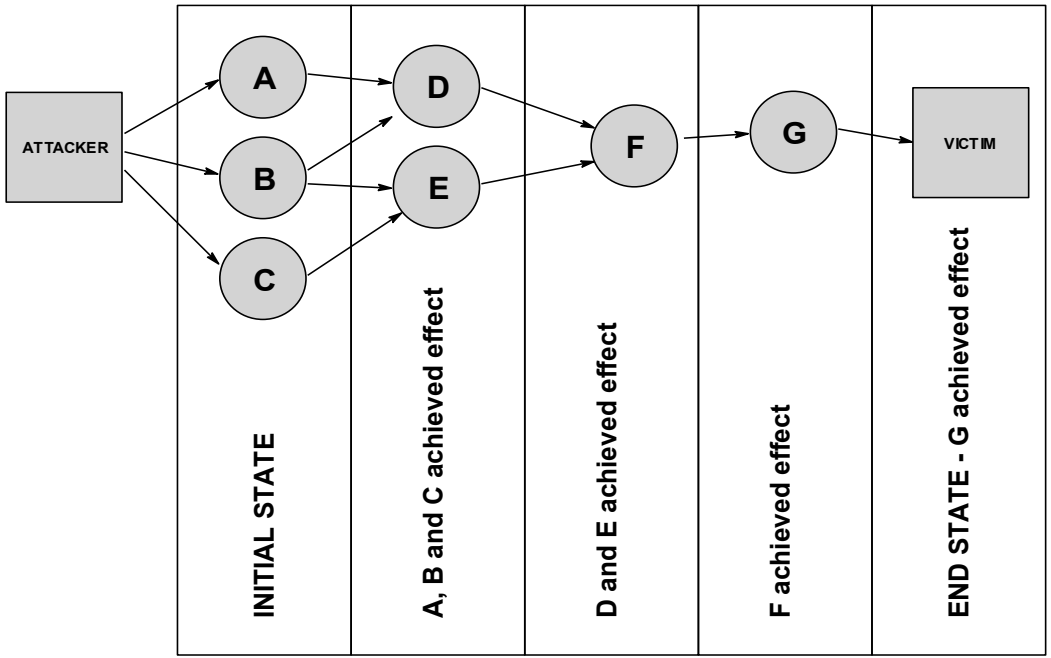


A Chained Compound IW Strategy

MODELLING COMPOUND ATTACKS

- A model for a complex compound attack is a directed graph, in which precedence relationships exist between component canonical attacks.
- The topology of this graph is dependent upon the structure of the compound attack.
- The overall success of any complex compound attack is measured by the end state of the victim. If the intended end state is not achieved, the strategy has failed.
- In terms of systematically constructing a compound information attack, the starting point is the end state of the victim, and the intermediate states the victim must transition between from its initial state.

State Transitions

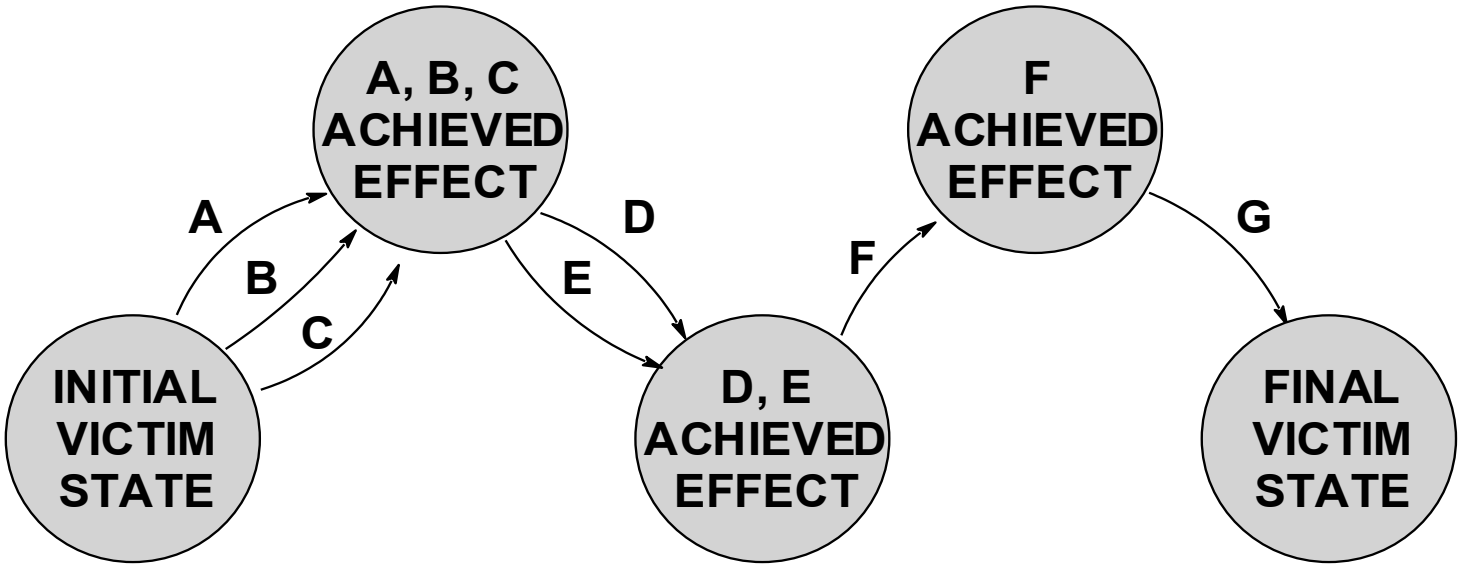


Victim State Transitions

STATE BASED MODELLING

- Alternate mappings for this modeling technique exist.
- A state based mapping is an alternative - attractive to users familiar with state transition diagrams, or project scheduling techniques such as PERT (Project Evaluation and Review Technique).
- In a state based representation, the graph comprises nodes which represent initial, intermediate and end states for the victim, and directed edges which represent the strategies required to effect a transition from a preceding state.
- Rather than searching for cut vertices in the directed graph, analysis requires that *bridges* be identified (Chartrand, 1977; Wilson, 1985).

State Based Representation



State Based Representation for Compound Strategies

Cut Vertices

- As compound information attacks have the properties of directed graphs, the behaviour of the cut vertex is of particular interest.
- A cut vertex is such a vertex, the removal of which partitions the graph into two smaller graphs (Chartrand, 1977; Wilson, 1985).
- Any attack, canonical or compound, which possesses the cut vertex property is a vulnerability within the overall compound information attack.
- The failure of this particular attack, or its defeat by the victim, results in the total failure of the whole attack.
- *Cut vertices are thus a critical vulnerability in compound information attacks.*

Robustness of Compound Strategies

- The attacker can assess the robustness of the attack at each state transition, by identifying whether the required attacks to effect that state transition have the cut vertex property, and thus represent a single point of failure for the compound attack.
- Robustness could be improved by executing two or more concurrent compound attacks, all of which effect the same end state in the victim.
- This is an application of the established reliability engineering technique of ‘parallel redundancy’ (Bazovsky, 1961).
- Example: 1944 Fortitude operation (Ministry of Defence, 2004; Ricklefs, 1996).

Defining a Metric for Robustness

- In defining a metric for calculating robustness we require a measure which can capture how robustness declines with the increasing number of cut vertices or bridges in a compound strategy.
- If we attribute some probability of failure to each of *N* cut vertices or bridges, then for equal probabilities, the probability of the compound strategy can be expressed as:

$$P_c[success] \quad = \quad (1 - P_i[failure])^N$$

- Where N is the number of cut vertices (or bridges) in the compound attack. This is *Lusser's product law*.

Generalising the Robustness Metric

- In a complex compound attack, the probabilities of failure associated with specific cut vertices or bridges may differ.
- Therefore the more generalised form applies:

$$P_c[success] = \prod_{i=1}^{N} (1 - P_i[failure])$$

- This model assumes no parallel redundancy in the graph, ie the loss of any cut vertex or bridge causes the whole strategy to fail.
- Where the compound strategy contains redundant paths, or dependencies exist between paths, then more general modelling techniques used in reliability engineering would be required.

Key Points

- Systematic analytical technique for modelling and analysing compound information attacks exist.
- Compound attacks are modelled as directed graphs, with precedence relationships where applicable.
- Discrete state transitions in the victim can be used as a measure of success.
- The concept of robustness in a compound attack is introduced, this being defined as a measure of how few component attacks in the compound strategy possess the cut vertex property.
- Future research is required to further explore techniques for the analysis of attacks *in progress*, techniques for modelling partial effects upon victims, and the effects of belief (false or true) in attackers and victims.

MONASH University

End Backup Slides

Understanding the Deception Pandemic

Dr Carlo Kopp, Fellow LSS, A/Fellow AIAA, SMIEEE, PEng
Faculty of Information Technology, Monash University, Clayton, 3800

GROUP OF EIGHT AUSTRALIA