The Phish in the Pond: Scam emails as Literature

Mark Blythe
Department of Computer Science
University of York, UK
mblythe@cs.york.ac.uk

John Clark
Department of Computer Science
University of York, UK

"Why, if a fish came to *me* and told me he was going on a journey, I should say, 'With what porpoise?'"

"Don't you mean 'purpose'?" said Alice.

"I mean what I say," the Mock Turtle replied, in an offended tone.

Alice in Wonderland.
Lewis Carroll

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
This paper considers phishing emails as literature. It reports a content analysis of a sample of phish taken from an online archive. The findings indicate that phishers are becoming better spellers and using more sophisticated visual aids such as logos and advertising images. The paper then considers phish as a literary form from two perspectives drawn from literary and critical theory: structuralism and psychoanalysis. It identifies the structural elements of phish and argues that we are currently vulnerable to them because the language of online security places such a heavy emphasis on individual responsibility for online security.

Keywords
Phish, security, structuralism, psychoanalysis

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous. See [3] for help using the ACM Classification system.

Introduction
Scam emails which attempt to lure people into revealing their banking details or visit websites which will install malicious software are so common and so old that there are now classics of the genre. Everyone who

has an email account will, sooner or later, receive a request for the temporary use of their bank account so that untold millions can be deposited following the death of some wealthy banking client. Very often such emails will be from Nigeria and indeed are known as "419" scams after the section number of the Nigerian penal code that deals with them.

The numerous guides to spotting these and other such "phish" agree that we must all follow simple rules:

- Never click on a link in an email
- Never respond to an email asking for confirmation of banking details
- Only use up to date virus protection, spam filters, operating systems and web browsers

Forms of address in emails such as "dear valued customer" are also often warned against. Poor spelling and grammar in communications from large organizations are also suggested as warning signs.

Use Experience and Critical Theory

Dourish et al note that detecting and deleting phish is a part of the more general practice of managing junk mail (Dourish et al 2004). The user experience of phish is a primarily literary one. An ongoing series of in depth interviews with blind users at York indicates very clearly that identifying phish is primarily a work of textual analysis. The application of literary theory to phish is then particular appropriate.

Content Analysis

MillerSmiles is an anti phishing website with a massive archive of reports. The following tables refer to 100 scams collected and archived for the period of 26th to the 31st of October 2009.

Category	Frequency
Security update	17
Invalid login	15
Security check	9
Account update	9
Update Records	5
Account suspension	4
Billing problem	4
Reactivate account	4
Security message alert	4
Security tips	4
Advertising	3
Confirm account	3
Unspecified	3
Changes notification	2
New message	2
Reset password	2
Warning	2
Work form home	2
Prize money	2
Different locations	1
Ebay user complaint	1
Racism	1
Unusual account activity	1

The most frequent phish were "security updates" followed by alerts caused by an invalid logins. Both are reasonably plausible stories as many systems do send real emails requesting changes of password following unsuccessul logins (e.g. JES). Unspecified "security checks" were also common along with general "account updates". There were four emails which purported to indicate that new security messages were waiting to be opened in a secure website. Each of these strategies draw on real security protocols in order to direct victims to phishing sites. Just four of the phishes were anything but banking security scams. One purported to be a complaint from an ebay user who had sent money and received no goods. Two were happy to announce a big lottery or prize win based on radom email selection. One was a call to click agreement with racist sentiments.

SPELLING.
Although poor spelling is often suggested as an indicator of a phishing attack, a preliminary content analysis suggested that some phish are better spelled than others. Just 11% contained three or more obvious spelling errors. Half contained one or two. The most common mistakes were missing or incorrectly used definite articles; mistakes with tense forms "after this will be completed" and incorrect forms of words "we advice you" rather than we advise you. But 38% contained no obvious errors at all. When asked what he would do if visited by a fish the mock turtle tells Alice it would enquire as to its porpoise. When visited by phish we also look to word slips, spelling mistakes and poor grammar are immediate give aways. But what happens when the phishers learn to present themselves in more convincing ways?

LOGOS
It has long been understood in advertising that there is tremendous value in a recognisable logo. Brand recognition is an important part of establishing trust. Logos have never been easier to copy and paste into documents and the use of visual aids such as these may lend credence to even poorly written phishes. Sixty four percent of the phishes contained a logo, some featured other visual aids such as colour schemes, photographs and graphically designed layout copied from official campaigns of the targetted company.

Literary Devices

The term "literary device" refers to particular techniques of writing. This includes technical devices such as the use of onomatapeia in comic books (boom, crash, pow); or broader and more complex patterns of organisation within a text like the use of symbolism and imagery in the novels of Virginia Woolf.

The first thing to notice about phish from a literary perspective is that they are all forms of pastiche. Pastiche is a form of imitation: the style and form of a particular author or a genre are drawn upon to create a new piece of writing. The vast majority of these phishes are pastiches of circular business letters from banks. Most of them are rather bad pastiches but a worrying minority are quite convincing. Not merely because they are spelled correctly and use the bank logos, they are also stylisticaly convincing pastiche. Because the form of writing being pastiched (or directly copied) is that of the business letter most of the devices are associated with formality.

The forms of address are courteous, sometimes exaggerated to the point of being courtly. Often these attempts at formality fail because of poor spelling or grammar but they can also fail by going too far "attention honourable beneficiary" or "Esteemed customer". The strategies draw on previous limitations of legitimate circulars. Mass mailouts mean that anonymous address is used for legitmate reasons "dear householder". The absolute impersonality of the email is stressed, it is merely part of a routine check. Similarly when threats are made they are rarely personally menacing, the threat is a matter of bureaucratic routine and the consequence of ignoring it is usually inconvenience rather than calamity.

Although the vast majority of the phishes pastiched business letters from banks another source for imitation was advertising. Some of the most convincing phish

took this form, adopting logos, formatting and images from plausible campaigns on security. Within the format of the genre imitated – the business letter, the advertising campaign, there are other literary devices drawn on to convey plausiblity. Some are very crude such as the invocation of a well known and trusted name "Microsoft" or a promise of upgrades to " the latest technology". But a minority are quite subtle and adopt a conversational and humerous tone.

With a subject line of "DO I EVER AGREE!!!!!" this phish appeals to patriotism and racism through (a sort of) humour:

> "How all business phones should be answered!
> 'GOOD MORNING, WELCOME TO CANADA
> Press '1' for English; and,
> Press '2' to disconnect until you learn to
> speak English.
> And remember there only two defining forces
> that have ever offered to die for YOU:
> Jesus Christ
> And our Canadian Soldiers.
> One died for your soul
> The others for your freedom.
> If you agree.....keep it going, if not
> simply hit the delete key"

Again, the form is a pastiche of an "email forward" which is to be passed amongst like-minded people.

An interestingly large proportion of these phishes employed forms of brevity "You have a new security message from HSBC. Click here. The Internet Banking Team at HSBC". Even briefer emails contained links which promised further information. Some of these stressed their own automation. Two were little more than lists of automatically generated text on when an email was sent and received with a link for the curious to find out more. These are interesting because rather than attempt to come up with a plausible reason for contact they rely on a blank appeal to curiosity.

Formalist / Structuralist Reading

Formalism is a branch of literary theory which originated in Russia in the early years of the twentieth century. Vladimir Propp, for instance, identified common elements in Russian folk tales which could be endlessly recombined to produce new stories. This work has been taken up to develop story engines (Braun 2004). Structuralism drew on mid twentieth century anthropology and its accounts of ritual and myth in European colonies (Easthope and McGowan 1992). Like formalism, structuralism sought to break narrative into its constituent elements. Almost before it began it was superseded by post-structuralism which argued that meaning is not constructed in discreet units. It is worth noting however that both terms are retrospective and nobody writing at the time declared themselves either "structuralist" or for that matter "post-structuralist" (Eagleton 2003). Nevertheless there is continuing interest in breaking complex things like narrative, or experience, into more tractable constituent forms and most books on media theory supply one list or another of the ingredients of story. What then might a structuralist analysis of phishing emails look like?

The constituent parts of a phishing email are something like:

- Interpellation
- Premise
- Instruction

INTERPELLATION

For the structuralist critic Louis Althusser "interpellation" is the process by which the state constitutes a subject. His most famous example is a policeman hailing a man on the street with the words "Hey, you there!" The man is unsure whether the policeman is addressing him or not but stops nevertheless and in so doing constitutes himself in a power relationship where he is subject to an ideological authority. Typical interpellations in phishing scams would be – dear valued customer, or client, or Sir or Madam. As previously noted phishers are subject to and exploit the same resources as junk mailers. They are also subject to the same limited forms of direct address to strangers. But as advertisers find it increasingly easy to use direct rather than anonymous address so too will phishers and other spammers. As other sources of contact from strangers develop it is likely that phishers will adapt to it – facebook messages, online game systems like Xbox, bulletin boards and so on will provide new opportunities for more effective interpellations

PREMISE

All phishing scams have a premise of one kind or another. Sometimes it is quite elaborate: a clerk at a bank discovers a way of accessing the account of a wealthy client who has died intestate, or quite simple: someone has tried to guess the password on your online bank account so it has been suspended. The main body of an email then will be exposition of one kind or another. There is a structural resemblance here to the joke, in that the set up is always a mis-direction. As with a joke the initial "feedline" raises expectations about how the rest of it is going to go. In a joke expectations are reversed with the punchline and the surprise causes laughter. In a phishing scam expectations appear to be fulfilled but results, at some later date, in fraud. As with jokes, if you've heard it before it will not be as effective. One of the main counter measures to phishers are programs of education. If people are educated about the way the scams work then they will not fall for them. However, the form of the security warning is now itself one of the most popular forms of phishing premise – there are new dangers, here are the ways to avoid them. Why this may be so will be returned to in the final section.

INSTRUCTION

The final basic constituent of a phishing attack is a call to action, some form of imperative command: confirm details, respond within forty eight hours and so on. For this to be successful the plot must be plausible to the victim: someone may have attempted to access their online bank account, there might be a problem with their PayPal account details, maybe their email really has been randomly selected in a lottery. For an instruction to be carried out there must either be belief or a willing suspension of disbelief. Primarily the calls to action play on fear. The genre then could be thought of as horror, but it isn't quite. It is not a desperate sort of fear, there is seldom real menace or the threat of physical harm. Perhaps suspense is a better way of framing this call. The instructions create suspense – what if it's real? What if there is a problem? The suspense is relieved when the link is followed or the bait is otherwise taken.

Decision making theory has been used to model user – phshing interaction into three stages:

"Construction of the perception of the situation; generation of possible actions to respond; generation of assessment criteria and choosing action"

(Dong et al 2008).

These stages correspond to the structural elements of the email content: the interpellation and premise construct a perception of the situation, the instruction generates possible actions. The model here is one of a user engaged in rational cognitions: "A user generates the criteria to evaluate the resulting gains and losses of possible actions, then evaluates those actions and chooses the best one." (Ibid). But what about irrational action and unconscious motivations? It is possible that such cognitive accounts of decision making may be supplemented with insights from psychoanalysis.

Psychoanalytic Reading

Psychoanalysis is now almost wholly ignored in departments of psychology. The success of cognitive psychology in both research and the treatment of mental illness has ensured that Freud and his factious successors have been consigned if not quite to the dustbin then at least to the literature department. In fields other than academia however, it has never gone away. The Public Relations industry was founded by Edward Bernays, the nephew of Sigmund Freud. Bernays explicitly drew on his Uncle's theories to make appeals to consumers which aimed not at their rational cognitions but their unconscious desires. Woody Allen famously spent half a lifetime on the couches of psychoanalysts and like many other famous cases, remains avowedly neurotic. However psychoanalysis has been and remains successful in exploiting our neuroses in order to persuade us to watch films or buy products or both.

Slavoj Zizek is a philosopher and critic who is a card carrying disciple of the French psychoanalyst Lacan and has written several books explaining Lacanian theoretical concepts through examples drawn from film and popular culture. Lacanian theory is based on a set of specialised terms and like all critical theory is often dismissed as jargon. Zizek defies those that reproach Lacan with being difficult with examples drawn from mass media and everyday experience. What might a Zizekian reading of phish look like?

Zizek is fond of a particularly gruesome story by Patricia Highsmith which is relevant here. "The Pond" is the story of a newly widowed woman who moves house with her young son. At first she loves her new home but at the bottom of the garden there is a dark pond clogged with strange weeds that she worries her child will fall into. She hires a firm to kill the weeds but the roots grow back almost immediately stronger than before. She tells her son not to go near the pond and warns that if he should ever fall in he must pull at the weeds to get to the side. But her son remains attracted to it and her fears become unbearable. She asks the pest control company to put more and stronger weed killer into the pond but when she gets off the phone she discovers that her son is missing. She finds him face down in the pond entangled in the weeds. After the funeral she returns to the pond and wades in to pull them out by hand. They now seem to be alive and the more she struggles against them, the more they drag her down into the dark water.

For Zizek the pond is "the sinthome" "the kernel of enjoyment that simultaneously attracts and repels us" (Zizek p133). Our fears do not simply appall us they also exert a strange fascination, we are uneasy about them, we return to them. We are uneasy about online security in just the way Highsmith's heroine is anxious about the pond. We return to it, we worry at it, and in

doing so sometimes open ourselves to the possibility of becoming ensnared. The successful phish both repels and attracts its victim. It keys into existing unease about online security and creates suspense. There could be no more perfect form for a phishing attack then, than a warning about a phishing attack with instructions on how to avoid it.

Literary theory is particularly sensitive to the relationship between form and content. The form here is a pastiche of the discourse of personal online security as espoused not only by banks but also those who sell protection. Here the focus is relentlessly on the individual, we must protect ourselves and if we don't we have nobody else to blame, certainly not the bank or the security firm. We must take some action or other and continually anticipate new kinds of attack. The standard warnings and tips demand eternal vigilance and constant updates. We are to be continuously afraid and endlessly fascinated; we must be cautious and yet quick to act. It is this impossible position which makes the language of security the perfect medium for fraud.

The emphasis on personal responsibility for online security is an ideological one. Although we are all well aware that we should update our anti-virus software, our firewalls, operating systems and web browsers we do not necessarily do it. Or at least not as often as we know we should. Like Highsmith's heroine we try but always fail to protect ourselves. And the sense of continuous anxiety which the ideology necessitates is the one which phishers are currently exploiting.

Conclusion

The content analysis indicated that phish are becoming harder to spot. Spelling and grammar are improving so the obvious give-aways may not serve as well in future. As social networking sites make it easier to gather personal information it is likely that phishers will begin to address us directly rather than as "valued customers". A structural analysis outlines the basic elements of a phishing attack: interpellation, premise and instruction. A psychoanalytic perspective indicates the complex aspects of user experience at play which continue to make us vulnerable.

References

[1] Braun, N. (2004). Storytelling and Conversation to Improve the Fun Factor in Software Applications. In M. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright, *Funology: From Usability to Enjoyment* (pp. 233-2243). Dordecht: Kluwer.

[2] Dong X., Clark J. A., Jacob J., Modelling User-Phishing Interaction. Human-System Interaction May 25-27, 2008, Kraków, Poland

[3] Dourish, P., Grinter, E., Delgado de la Flor, J., and Joseph, M. 2004. Security in the wild: user strategies for managing security as an everyday, practical problem. Personal Ubiquitous Comput. 8, 6 (Nov. 2004), 391-401

[4] Eagleton T. (2003) After Theory. London, Penguin Books

[5] Easthope, A., & McGowan, K. (1992). *A Critical and Cultural Theory Reader.* Milton Keynes: Open University Press.

[6] Propp, V. (1968). *Morphology of the Folk Tale.* Texas: University of Texas Press.

[7] Zizek S., (1992) Looking Awry: an introduction to Jacques Lacan through popular culture. October Books