Mainstream Advertising on Rogue Websites in Malaysia:

A Comparison of Local and Foreign Content

Paul A. Watters

Massey University

Executive Summary

A number of studies have recently investigated the role played by mainstream internet advertising in supporting the revenue of rogue websites (Taplin, 2013). Such advertising by household names – including multinational corporations, governments and charities - generates enormous profit margins for operators of these websites, and present an ongoing threat to the viability of Malaysia's creative industries.

However, a recent study by Watters (2014a) indicated users were much more likely to be exposed to "high risk" advertising on such sites, relative to mainstream ads. Australia (Watters, 2014a), Singapore (Watters, 2013a), Canada (Watters, 2013b), and New Zealand (Watters, 2014c) all had mainstream prevalence rates of 1-10%, while high risk ads had prevalence rates of 90-99%. High risk ads are those which have the potential to cause harm to users, and include pornography, gambling, malware and scams.

These studies all investigated web pages that were sampled from Google's ad transparency report for movies and TV shows or music downloads, having been verified as being in breach of the Digital Millennium Copyright Act (DMCA)[1]. However, the Google report is heavily biased towards Hollywood TV/movies and music in English, so another study (Watters, 2014d) investigated Hollywood content in Taiwan, where the sites were presented in Chinese. It was found that mainstream advertising was much more prevalent for locally-developed sites with Hollywood content, compared to viewing Hollywood content sites in the other countries examined. 61% of ads were Mainstream, while 39% were High Risk for local content. A follow-up study in Hong Kong (Watters, 2014e) investigated whether Mainstream advertising would also be more prevalent not just for sites written in the local language, but also promoting local content (eg, a Chinese language website providing links to Chinese language titles. For local content, 61.36% of movie and TV ads were Mainstream, while 38.64% were High Risk. In contrast, 3.84% of Hollywood ads were Mainstream, while 96.16% were High Risk.

In this study, we directly measured prevalence rates for High Risk versus Mainstream ads for local versus Hollywood content for movies/TV in Malaysia. A sample of expert-identified rogue sites hosting local content was identified, and all ad banners comprising part of the sample were downloaded and identified, along with the ad network serving each banner. A comparable sample was taken from Google's ad transparency report. It was found that for local Malay content, 72.1% of movie and TV ads were Mainstream, while 27.9% were High Risk. In contrast, 8.24% of Hollywood ads were Mainstream, while 91.76% were High Risk. Like Hong Kong and Taiwan, this suggests that

[1] Note that the URLs for each of the sites for which there was an upheld complaint were accessed through the Chilling Effects database (http://www.chillingeffects.org/).

Mainstream advertisers in Asia are being drawn to local language sites, whereas Mainstream advertising rates for Hollywood titles are similar to other countries.

In summary, local content sites were many times more likely to be displaying mainstream ads when compared to Hollywood content sites. The levels of mainstream advertising were almost identical to Taiwan and Hong Kong for local content, and were similar to Canada, Singapore, Australia and New Zealand for Hollywood content.

The policy implications of this result and future research directions, including methodology enhancements, are discussed.

Keywords

Infringing content, internet advertising, Digital Millenium Copyright Act (DMCA), internet safety.

Definitions

Internet Advertising. Ads are typically placed as "banners" on a website, which direct a user to another site when clicked. The contents of the ad are similar to a highway billboard, except that they can incorporate interactive elements such as animation. Ads on the same page are often rotated through a predetermined or random sequence, depending on the advertising plan that an advertiser has subscribed to. While some sites host and manage their own banners, most often, these are managed by a third-party advertising network. These ad networks act as an intermediary between an advertiser and many hundreds, thousands or millions of sites, allowing an advertiser to increase their reach to potential consumers while only dealing with a single agency. Advertisers typically operate either a "pay per impression" or "pay per click" model, billing an advertiser every time a user views or clicks on a banner ad respectively.

Mainstream Advertising. Mainstream ads are those placed by legitimate businesses that operate within the formal economy. Such businesses operate through a corporate structure and offer goods or services which fall outside the black market, grey market or underground economy.

High-Risk Advertising. High-Risk ads are those promoting goods or services which fall outside the legitimate economy or white market, may be illegal or restricted within certain jurisdictions but not others, or may be fake or counterfeit. Examples include the sex industry, gambling and suspicious software/malware, such as anti-virus software which actually installs a Trojan Horse on a user's system. Many of the ads are likely to fall into scam categories described by Stabek et al (2009).

Advertising Network. Ad networks facilitate the placement of an advertiser's ads on numerous websites according to a specific revenue model. Ad networks specialise in anticipating consumer's needs and wants by building up profiles of users who click most frequently on certain ad categories on certain page themes, which can lead to more targeted, personalised, and relevant advertising. For the purposes of this paper, sites that host advertising on behalf of external / third-party advertisers are also grouped under this category, even if they only provide banners on sites within their own domain. For example, isohunt.com provides their own ad network exclusively for their own site, and not to other sites; they also host banners from other ad networks.

Internet Advertiser. A business, government, association or individual that desires to sell goods or services, or provide information to, a target group of consumers. Internet advertising competes with traditional advertising for marketing budgets. Malaysia's online advertising market was valued at RM14.9 in 1H2007[2] and is growing at a rate of 15.9% pa, suggesting current revenue of around RM188.99[3] in 2014.

Rogue Site. A website which provides an index and search capability for torrents of infringing content, a "file locker" site which provides hosting for such material (through file sharing or direct download), or a "link site" which provides direct links to content on third party sites. The primary motivation for users visiting these websites is to access infringing content. These sites can all use advertising as either primary or secondary sources of income.

Digital Millenium Copyright Act (DMCA). The DMCA provides ISPs with indemnity against liability for copyright infringement, provided that they agree to co-operate in "takedowns" of material which is alleged to be infringing, typically after being notified by a rightsholders. Google provides a report of requests that they have received and actioned on behalf of rightsholders in order to provide transparency to their users.

Introduction

Online advertising has a 20 year long history (Medoff, 2000), progressing from simple ad banners displayed on a fixed rotation schedule, through to personalised, behavioural advertising networks, which use profiles of individual users to present the most "relevant" advertisements (McStay, 2011). Such technologies make extensive use of "tracking cookies" (Watters, 2012) and the linkages between advertising networks and cookies have recently been monitored and explored for the most popular websites in Australia (Herps et al, 2014). The most interesting result from this study was that the number of cookies stored on a user's computer from any of the Top 50 most-visited sites for Australians ranged between 0 and 86. The sophistication and the extent to which user behaviour is tracked and experiences customised is only going to increase over time, as is the overall volume of advertising. Indeed, in 2012, online advertising spending in the US reached US$39.6b, exceeding the amount spent on traditional print advertising for the first time (eMarketer, 2012), and predicted to reach $44.74b for 2013 and $51.01b in 2014.

Furthermore, some companies are in a unique position to know "everything" about their customers. Google, for example, has the capacity to monitor almost all of the world's information, including personal emails, YouTube movies, Android phones, news services, images, shopping, blogs and so on (Cleland, 2013). Through its acquisition of Doubleclick, Google controlled an estimated 69% of the online advertising market (Browser Media, 2008), however, the rise of social media advertising (especially through Facebook) has seen this reduce to 56% (Womack, 2013). Clearly, there is a potential confluence of capability and opportunity to maximise the number of "eyeballs" exposed to online ads.

[2] http://www.skmm.gov.my/skmmgovmy/files/attachments/Ad_Dev_Malaysia.pdf
[3] http://read.pwc.com/i/179445/12

What are the implications of this massive rise in advertising expenditure, which coincides with an increased ability for online advertising networks to be able to best "place" ads to suit specific customers? One particular type of website – those associated with file sharing of infringing content – appears to have wholeheartedly embraced advertising. Indeed, advertising revenues provide the commercial motivation for criminal syndicates to operate such 'rogue' web sites. While the connection between film and television piracy and organised crime has been explored elsewhere, in terms of direct revenues (Treverton et al, 2009), there has been far less publicity about the advertising revenues generated from sites that appear to offer infringing content for free, or at least, offer torrents that enable users to download such material. Certainly, the links between the underground economy and the internet have been criticised for facilitating sexual exploitation and human trafficking through organised crime – in the classic paper in this field, Hughes (2000) highlighted how global advertising and marketing of prostitution have led to increases in volume globally. Furthermore, Hughes identified that a *lack of regulation of internet advertising* was the key policy failure in preventing harm to women and children.

The Pirate Bay is one of the most popular sites for providing torrents to infringing content, and has been the subject of criminal proceedings against its operators in Sweden. In the 2009 trial of its operators, their expenses were estimated to be US$110,000 p.a (Olsson, 2006; Kuprianko, 2009), with advertising revenues in the order of US$1.4m p.a (Sundberg, 2009) – in other words, an extremely profitable business with gross margins of 1272%! A recent study (Detica, 2012) indicated that there are six different business models operating within the pirate site marketplace, ranging from advertisement and donation funding, through to subscriptions and freemium sites, where subscribers can gain faster access to illicit content by paying a subscription fee. 83% of the sites in that study operated using a central website. Selling advertising on file locker and torrent search sites is the major source of revenue for such sites. The Pirate Bay, for example, regularly features in the Top 50 sites accessed by Malaysia residents (as computed by alexa.com) , and so it is a potentially attractive space for advertisers and ad networks, since the number of potential "eyeballs" is very high. Other rogue sites with high Alexa rankings include Kickass torrents (rank 103) and Torrentz (rank 153)[4].

Maximising "eyeballs" leads to clicking, which drives revenue for the ad networks (if they operate a Pay Per Click revenue model), and sales for the advertisers. A key question for advertisers and ad networks is the extent to which they wish to be associated with this type of activity; indeed, due to the complex algorithms which decide which ads to display to which users, advertisers may not be aware of every site that their ads are being displayed on.

Being able to quantify the scale of advertising on these sites is important, since informing and making advertisers aware of the integrity of the sites on which their ads are being displayed can then be undertaken. Advertisers will thus be able to make more informed choices about their use of online advertising networks (the companies who provide aggregation of space on web sites) who are supporting piracy by selling ad space on torrent and file locker sites. A recent set of best practice

[4] http://au.ibtimes.com/articles/533033/20140106/pirate-bay-popular-torrent-site-top-10.htm#.Uvk8g_mSw3I

guidelines for ad networks to address piracy and counterfeiting have recently been released[5], and early indications are that most of the world's major web companies will participate[6].

There have been few systematic studies investigating the relationship between piracy and advertising, and most have been concerned with the impact of interventions to reduce piracy. For example, Sheehan et al (submitted) identified that increasing the perception of legal risk for college students was most likely to influence downloading behaviour, while Gopal et al (2009) weighed up the ethical predispositions of downloaders and their beliefs in justice and law to the money potentially saved by downloading infringing content. Indeed, it is this appeal to justice as the primary virtue of social behaviour (Rawls, 1999) that may concern ethical advertisers if their advertising expenditure was being used to fund illicit activities.

A number of studies[7] have recently examined the relationship between piracy sites and online advertising networks (Taplin, 2013). The USC report provides a method for revealing the advertisers whose ads are most likely to be served up on these sites, which may be occurring without the direct knowledge of the advertiser. While the objectives of USC research are significant, the monthly rankings of the "top ten" advertising networks responsible for placing the most ads on web sites that support infringing content are surprisingly variable – Google, for example, was ranked at #2 in January 2013, but did not appear at all in the February and March 2013 lists at all. One interpretation of the result could be that the January report achieved its goal of sensitising advertising networks, and that Google subsequently withdrew from placing ads on those sites. Alternatively, the variation could be due to biases inherent in studies using an observational methodology, including selection bias, information bias and recall bias. The lack of detail in how measures like the "top 500" sites prevent the study results from being directly replicated, which would be the standard required for peer review by other researchers. By not providing this level of detail, the credibility of the USC report may be called into question by the very vocal critics of any research in the anti-piracy field.

The first peer reviewed paper in this field was published by Watters (2014a). That study outlined a fully replicable algorithm for sampling rogue sites to provide a much clearer view of advertising network behaviour in different countries, jurisdictions, languages etc. The major difference between this study and the USC study was that it examined all advertisements, not just the Mainstream ones. In doing so, Watters was able to establish the relative proportion and prevalence of Mainstream ads versus High Risk ads. High risk ads are those which have the potential to cause harm to users, and include pornography, gambling, malware and scams. After examining the ads being served to Australians in that study, it was found that 99% of the ads from the "top 500" sites were High-Risk, while only 1% were Mainstream. Subsequent studies measured the prevalence rates for Holly movies/TV in Singapore, Canada, and New Zealand[8], finding that prevalence rates for mainstream varied between 1-10%, while high risk ads had prevalence rates of 90-99%.

[5] http://2013ippractices.com/bestpracticesguidelinesforadnetworkstoaddresspiracyandcounterfeiting.html
[6] http://torrentfreak.com/tech-giants-sign-deal-to-ban-advertising-on-pirate-websites-130715/
[7] Not peer-reviewed

These studies all investigated web pages that were sampled from Google's ad transparency report for movies and TV shows or music downloads, having been verified as being in breach of the Digital Millenium Copyright Act (DMCA). However, the Google report is heavily biased towards Hollywood TV/movies and music, so another study (Watters, 2014d) investigated "local" content in Taiwan (Chinese language), and compared the prevalence rates for high risk versus mainstream ads with Hollywood content. It was found that mainstream advertising was much more prevalent for local content sites compared to viewing Hollywood content sites in the other countries examined. 61% of ads were Mainstream, while 39% were High Risk for local content. However, no direct comparison was undertaken to see whether Taiwanese viewers were more likely to see mainstream content on local content sites versus Hollywood sites. A subsequent study (Watters, 2014e) undertook such a comparison, by measuring the prevalence rates for High Risk versus Mainstream ads for local versus Hollywood content for movies/TV in Hong Kong, finding that Mainstream advertising for local content was far more prevalent than for Hollywood content.

This study aims to replicate the Hong Kong study; prevalence rates for local Malay content will be compared to Hollywood titles. It is predicted that, for local content, the prevalence rates will be similar to Taiwan and Hong Kong, but for Hollywood content, the prevalence rates will be similar to Australia, New Zealand, Canada and Singapore.

Methods

The main goal of the methodology is to identify the advertising networks and advertisers from a sample of DMCA complaints, which have been ranked in terms of the number of complaints upheld by Google (through their Transparency Report)[9], or from a list of expert-identified rogue sites. The DMCA complaints typically relate to the availability of search results for a wide range of potentially infringing content; by only selecting the most complained about and subsequently upheld complaints as assessed by a third-party (Google), the results should be robust against criticisms that there is no proof that the sites in question were hosting torrents of infringing content or infringing content directly, in the case of a file locker site. The methodology operates by downloading each page from the "top 500" complaints submitted to Google within the previous month, ordered by the number of upheld complaints. Since each DMCA notice can contain many thousands of individual URLs, a sampling procedure can be used to identify a representative subset of URLs, and the advertisements on each page can be downloaded along with their metadata. In the case of simple banner ads, it is then relatively easy to identify the advertisers concerned; in the case of each distinct advertisement, a rule can be generated using SQL or similar to identify all advertisements with the same metadata. However, some advertising networks use JavaScript obfuscation and a series of redirects to obscure the ultimate destination for the advertising banner; in this case, manual inspection must be performed, in the absence of a general purpose image/logo recognition system. The overall prevalence of a particular advertiser on each network can be then be computed and ordered by frequency. An example obfuscated URL is shown below, from adcash.com:

```
var ct_SuLoaded = 1; var ct_SuUrl='http://' + 'www.adc' + 'ash.com' +
'/script/pop_packcpm.php?k=52e9e900315f7850142.1482334&h=e0004318879f95130e
```

[9] Google.com.my is ranked as the #4 site for Malaysians to visit; http://www.alexa.com/siteinfo/google.com.my

```
6cf975c6b761523a69e256&id=0&ban=850142&r=128995&ref=&data=&subid=';     var
ct_nSuUrl='http://'       +       'www.adc'       +       'ash.com'       +
'/script/pop_packcpm.php?k=52e9e900315f7850142.1482334&h=e0004318879f95130e
6cf975c6b761523a69e256&id=0&ban=850142&r=128995&ref=&data=&subid=&new=1';
// ct_SuUrl += '&sw=' + screen.width + '&sh=' + screen.height; // we are
hiring!  say  hello  at  \x72\x6f\x6f\x74@adcash.com  var  evaI  =  eval;var
ct_siteunder         =         true;var         ct_tag         =         1;
evaI(function(p,a,c,k,e,r){e=function(c){return(c35?String.fromCharCode(c+2
9):c.toString(36))};if(!''.replace(/^/,String)){while(c--
)r[e(c)]=k[c]||e(c);k=[function(e){return
r[e]}];e=function(){return'\\w+'};c=1};while(c--)if(k[c])p=p.replace(new
RegExp('\\b'+e(c)+'\\b','g'),k[c]);return  p}('(6(){h  G=l.n.m().o("2v");h
13=l.n.m().o("2S");h         w=l.n.m().o("2u");h         z=l.n.m().o("35");h
1p=l.n.m().o("2A");2(w!=-1){h              I=1Q(l.n.m().1A(w+7))}2(z!=-1){h
2n=1Q(l.n.m().1A(z+7))}2(1p!=-1){h   2t=1Q(l.n.m().1A(1p+4))}2(G==-1&&13==-
1&&w==-1&&z==-1){h     13=l.n.m().o(".36")}h     S=i;h     5=W;h     F=W;6
k(a,b,c){2(a){2(a.2s){a.2s(b,c,t);K i}j 2(1D.1F){a.1F("2m"+b,c);K i}}K t}6
1c(a,b,c){2(a){2(a.2l){a.2l(b,c,t);K i}j 2(1D.2T){a.1F("2m"+b,c);K i}}K t}6
2k(){S=i}6   2h(){S=t}6   1u(){g.2w("2z:g.J()",W,"2B:2E;2F:2G;2H:1;2J:1")}6
u(){2(17)K;17=i;h                                  1H=l.n.m().o("2Z");2(G!=-1)h
1e=l.n.m().20(G+8);2(1e>=1&&l.n.m().20(G+9)>=0)1e=10;2(G!=-
1&&1e>3){q=g.x("","2D","1i=1,Z=1,1l=1,1n=1,1m=1,19=1,1k=1,1j=1,M="+N+",O="+
P)}j 2(w!=-1){q=g.x("","1h","1i=1,Z=1,1l=1,1n=1,1m=1,19=1,1k=1,1j=0,M="+(N-
10)+",O="+(P-1f))}j                                                    2(z!=-
1){q=g.x("","1h","1i=1,Z=1,1l=1,1n=1,1m=1,19=1,1k=1,1j=0,M="+(N-
10)+",O="+(P-1f))}j        2(1H!=-1){q=g.x("","1h","M=1s,O=1s");q.2y(1a.1B-
1f,1a.2f-1f);q.2C(0,0)}j                                                2(13!=-
1){q=g.x("","1h","1i=0,Z=1,1l=1,1n=1,1m=1,19=1,1k=1,1j=1,M="+N+",O="+P+",1b
=0,1E=0");q.E();s(6(){q.E();g.r()},1J);s(6(){q.E();g.r()},1J)}j
2(G){q=g.x(1x,"1h","1i=1,Z=1,1l=1,1n=1,1m=1,19=1,1k=1,1j=1,M="+N+",O="+P+",
1b=0,1E=0")}2(2L){2(13!=-1){2(q.E)q.E();q.2M.g.r()}2(w!=-1){g.E()}j
2(q.E){q.E()}2((w!=-1||13!=-1||1H!=-
1)&&g.r){g.r();s(6(){g.r()},1J);s(6(){g.r()},2e);s(6(){g.r()},2c);s(6(){g.r
()},1s);s(6(){g.r()},2U);s(6(){g.r()},2V);s(6(){g.r()},2W)}2(G!=-1&&1e>3){h
L=t;h         p=g.x("16:12");p.r();p.J();L=i;s(6(){2(L&&!S){h         L=t;h
p=g.x("16:12");p.r();p.J();L=i}},2e);s(6(){2(L&&!S)1u()},2c);s(6(){2(L&&!S)
1u()},1s);s(6(){2(L&&!S)1u()},3x)}}2(q){q.Z=1x}}6         1d(a,b){2(f.1y){h
c=f.1y.2a("A");1r(h                    d=0;d=31||(I>=30&&l.n.m().o("2i")==-
1)){k(f,"T",6(e){2(e.D==0&&!v&&!1X){u();v=i}j
2(e.D==0&&!v&&1X){5=f.Q(\'a\');5.X="2j";5.1V=1x;f.U.R(5);h
e=f.1N("1I");e.1C("B",i,i,g,0,0,0,0,0,i,t,t,i,0,W);5.1K(e);5.11.18(5);v=i}}
)}j
2(I>=30&&l.n.m().o("2i")>=0){k(f,"B",6(e){2(e.D==0&&!v){f.3b.3c(3d.3e);u();
f.3f();v=i}})}j
2(I>=25){k(f,"T",6(e){2(e.D==0)u()});k(f,"V",6(e){2(e.D==0){2(!v){5=f.Q(\'a
\');5.X="3g";5.1V=\'3h:3i/3j;3k=3l-8,\'+3m(\'<2o\'+\'2p>g.J()\');f.U.R(5);h
e=f.1N("1I");e.1C("B",i,i,g,0,0,0,0,0,i,t,t,i,0,W);5.1K(e);5.11.18(5);v=i}}
})}j{2(I<=24){k(f,"T",6(e){2(e.D==0)u()});k(f,"V",6(e){2(e.D==0)2(!v){2(z!=
-1&&w==-1&&l.n.m().o("H")!=-
1){}j{g.x(\'\',\'\').J()}v=i}})}j{2(2q){1v(1z)}j{k(f,"2r",6(e){C=e.C;2(e.1w
)e.C=e.1w;2(C.1O!=\'1P\'){g.1T=C}});k(f,"T",1o)}}}}j                   2(z!=-
1){2(2n<3t){k(f,"T",6(e){2(e.D==0)u()});k(f,"V",6(e){2(e.D==0)2(!v){2(z!=-
1&&w==-1&&l.n.m().o("H")!=-
```

```
1){}j{g.x(\'\',\'\').J()}v=i}})}j{2(2q){1v(1z)}j{k(f,"2r",6(e){C=e.C;2(e.1w
)e.C=e.1w;2(C.1O!=\'1P\'){g.1T=C}});k(f,"T",1o)}}}j{f.3u=u;2(k(1D,"1t",6(){
1d()})==t){k(f,"1t",6(){1d()})};k(f,"B",6(){u()});1d();s(1d,3v)}g.s(6(){2((
I>=14&&I<=29&&l.n.m().o("H")==-1)||(z!=-1&&l.n.m().o("2u")==-
1&&l.n.m().o("H")==-1)||(1p!=-1&&2t>=15&&l.n.m().o("H")==-
1)){5=f.Q(\'a\');5.X="2j";5.1V=1x;f.U.R(5);h
1Y=f.1N("1I");1Y.1C("B",i,i,g,0,0,0,0,0,i,t,t,i,0,W);5.1K(1Y);5.11.18(5);17
=i;v=i}},2K)})()',62,220,'||if|||ct_gli|function|||||||||document|window|va
r|true|else|ct_addListener|navigator|toLowerCase|userAgent|indexOf||ct_pop|
focus|setTimeout|false|ct_showPopupWindow|ct_shown2|ct_agentchrome|open|sty
le|ct_agentsafari||click|target|button|blur|ct_ylz|ct_agentff|mobile|ct_age
ntchrome_version|close|return|pcycle|width|ct_popW|height|ct_popH|createEle
ment|appendChild|ct_focus|mousedown|body|mouseup|null|id|w2|location||paren
tNode|blank|ct_agentmsie|||about|ct_shown|removeChild|resizable|screen|top|
ct_removeListener|ct_hookLinkTags|ff_version|50|body_checking|_blank|toolba
r|menubar|copyhistory|directories|scrollbars|status|ct_showPopupWindow2|ct_
agentopr|length|for|350|load|ct_modal|body_check|srcElement|ct_nSuUrl|all|c
t_showPopupWindow3|substring|availWidth|initMouseEvent|this|left|attachEven
t|getElementsByTagName|agentopera|MouseEvents|100|dispatchEvent|iframe_mous
eup2|ct_klq|createEvent|className|norecord|parseFloat|iframe_load|callback_
function|hoverElem|display|href|iframe_mouseup|ct_tabunder|ct_e|contentDocu
ment|charAt|px|src|frame6164637368|||iframe|ct_removeDocumentEvent|IFRAME||
tags|none|250|submit|150|availHeight|25px|ct_setBlur|windows|a616463736899|
ct_setFocus|removeEventListener|on|ct_agentsafari_version|scr|ipt|ct_tag|mo
useover|addEventListener|ct_agentopr_version|chrome|firefox|showModalDialog
|pageX|resizeTo|javascript|opr|dialogtop|moveTo|adc|971009|dialogleft|99711
5104|dialogWidth|79|dialogHeight|142|ct_siteunder|opener|form|form141144143
141163150|141144143141163150|input|type|msie|detachEvent|750|850|1850|clear
Interval|setInterval|opera|||1024|768|frameborder|safari|net|hidden|backgro
undColor|transparent|opacity|documentElement|webkitRequestFullscreen|Elemen
t|ALLOW_KEYBOARD_INPUT|webkitCancelFullScreen|a6164637368|data|text|html|ch
arset|utf|encodeURI|cursor|pointer|position|absolute|zIndex|999|536|onclick
|2000|pageY|950'.split('|'),0,{}));// tbu6 + IE11
```


Furthermore, it may be of interest to separate out “Mainstream” advertisements as opposed to “High-Risk” advertising, since the Annenberg reports indicate a flight by Mainstream advertising this year from sites that host infringing content. Advertisers who may otherwise be unable to place their ads on a Mainstream site can then take advantage of increasing “eyeballs” by occupying display space. Results are thus reported for the High-Risk and Mainstream categories, with the former including categories such as:

- Sex Industry, which includes adverts for:
    - Penis length extension medication
    - Fake personal/dating sites
    - Pornography of various kinds
    - Dating and “foreign bride” sites
- Online Gambling
- Malware, including

    - Fake software incorporating Trojan horse malware (numerous alerts were raised by anti-virus software during the data collection process due to "drive by downloads" of malware). A Trojan horse is a piece of software that hides a malicious payload inside what appears to be an otherwise legitimate download.
    - Fake anti-virus or anti-scamware
    - Suspicious software such as fake video codecs or video players that replicate existing functions within Microsoft Windows. The purpose of such downloads is unclear, although it is possible that they could host Trojans or provide backdoor access to systems.
- Scams, as defined by Stabek et al (2010), such as:
    - Premium rate SMS scams
    - Fake competitions where no prizes are offered
    - Investment scams
    - Employment scams

The algorithm works as follows:

1. A data collection system is installed physically or logically to attract advertising for a specific geographical/country segment. For this study, Malaysia was selected.
2. The current Google Transparency Report[10] is downloaded, which lists all of the DMCA requests for a specific time period[11]. This list provides one means of identifying sites involved in sharing pirated material. In parallel, a list of expert-identified rogue sites targeting Malaysia local content was compiled.
3. The dataset is sorted by the number of URLs removed, retaining the "top 500" DMCA requests (the request list) by complaint category. For this study, the complaint categories were movies and TV shows; other complaint categories such as pirated software, adult material, music etc were excluded.
4. For each report in the request list first 10 URLs are extracted as a representative sample of all of the URLs contained within the report. This gives a total of 5,000 web pages to be downloaded (the sample) for Hollywood content. For local content, 36 sites were downloaded, giving a total of 360 local pages.
5. Each of the 5,360 web pages in the sample is downloaded, and a screenshot is taken, showing the ads being served. Note that for technical reasons, pop-up ads are not captured.
6. For each web page in the sample, the code blocks that contain advertising are parsed and extracted. This can be achieved by matching against the Easy List[12] (used by Adblock Plus for filtering), for known URL patterns and hostnames of advertisers. Some pages in the sample will have no ads, while others will have multiple ads.
7. For each advertising code block, the domain of the advertising network being used is identified, by stripping extraneous code and links from the code block, and counting the frequency of appearance of each ad network domain. If an ad network has fewer than 5

[10] https://www.google.com/transparencyreport/removals/copyright/data/
[11] The DMCA list for March 2014 was used in this analysis
[12] http://easylist.adblockplus.org/en/

occurrences, the items are discarded. The rationale for exclusion is that errors in coding, extraneous links etc can result in false positives being included in the list.

8. For each identified advertisement, an attempt is made to identify the actual advertiser, by analysing metadata, following the link and extracting the domain of the actual advertiser, or through visual inspection. A list of all identified advertisers is then generated.

Results – Hollywood Movies/TV

Appendix A contains a list of the DMCA notices identified in Step 3, including TV and movies from major Hollywood studios such as Fox, Warner Bros etc during March 2014. From the 5,000 pages analysed in Step 4, a total of 4,214 advertising items and 2,050 visible ads were identified in Step 6[13]. Postprocessing of the identified domains were performed to ensure that all ad blocks were correctly identified, for example, by removing port numbers that were included as part of a URL by using a regex filter. 242 unique domains for advertising networks were identified, indicating an average 8.47 visible ads per network in the sample (keeping in the mind that the distribution – shown in Table 1's Top 10 advertising networks - is highly non-uniform). Appendix B contains the complete list of advertising networks detected. Note that no merging of distinct services was performed, eg, the several domains of The Pirate Bay were not aggregated, to preserve the literal domains as observed. Also, where a domain appears within an ad block, this is a technical definition as per the methodology in Steps 6 and 7, ie, if the site or known ad URL appears in the block, then it will be counted. This could include Facebook social plugins, for example, rather than Facebook ads.

The analysis is presented by reviewing the High Risk ads first, followed by the Mainstream ads.

Table 1. Frequency Analysis by Advertising Network – Top 10 [14]

Advertising Network	Frequency	% of Ads
adcash.com	802	37.78%
adexprt.com	402	18.94%
depositfiles.com	213	10.03%
disqus.com	122	5.75%
filestube.to	120	5.65%
ismcorp.vn	99	4.66%
onclickads.net	97	4.57%
propellerads.com	92	4.33%
uasdel.com	92	4.33%
xtendmedia.com	84	3.96%

[13] Advertising items include any scripts, images, spacers etc being referenced from an Adblock domain, in addition to visible ads

[14] Note that some ad networks like isohunt.com and sumotorrent.com do not display their ads outside their own domain; they are ranked highly because of the high number of DMCA complaints against their site.

High-Risk Advertising – Top 10 (Hollywood Movies/TV)

Table 2 contains a summary of the results from the Top 10 ad networks. There were 2,123 advertising items in this sub-sample. Each of these advertisements was downloaded, visually inspected and categorised. The results indicate that malware, scams (including employment, investment and SMS premium rate), and the sex industry were the most popular distinct advertising types in Malaysia for the Top 10 networks.

An example of malware downloaded is provided by the advertising link **http://isohunt.com/a/adclick.php?bannerid=493&zoneid=&source=btDetails-banner&dest=http%3A%2F%2Flp.ncdownloader.com%2Fexact%2F%3Fq%3DCannonball+Run+II.+1984**. Upon visiting this page, a download is initiated to the user's computer containing the file **Cannonball Run II. 1984.exe** which is only 292K in size – much smaller than a typical video file of at least 700M. Running this file through the online scanner virscan.org – which analyses suspicious files using 36 different products – the file is verified as ADWARE/Adware.Gen (http://v.virscan.org/ADWARE/Adware.Gen.html) by AntiVir 8.2.10.202 and as Adware.Downware.1166 by ClamAV (http://v.virscan.org/Adware.Downware.1166.html). A review of the other known filenames associated with this malware indicates a typical strategy of associating a desirable filename with the malicious code, ie, using a filename that users desiring to download infringing content will click on, including **Mortal Kombat - Komplete Edition Crack (2013) Download.exe** and **Transformers 3 - Dark of the Moon (2011) [1080p].exe**.

Table 2. High-Risk ad type frequencies by network

Ad Network	Ad URLs	Distinct Ad URLs	Sex Industry	Malware	Down loading	Gaming / Gambling	Scams	TOTAL
propellerads.com	802	107		34				34
filestube.to	402	80		50				50
adcash.com	213	31	2	3		1		6
disqus.com	122	2						0
xtendmedia.com	120	2						0
depositfiles.com	99	9		2	3			5
adexprt.com	97	43	5	26		1		32
onclickads.net	92	4		58			1	59
uasdel.com	92	11						0
ismcorp.vn	84	28						0

TOTAL	2123	317	7	173	3	2	1	186

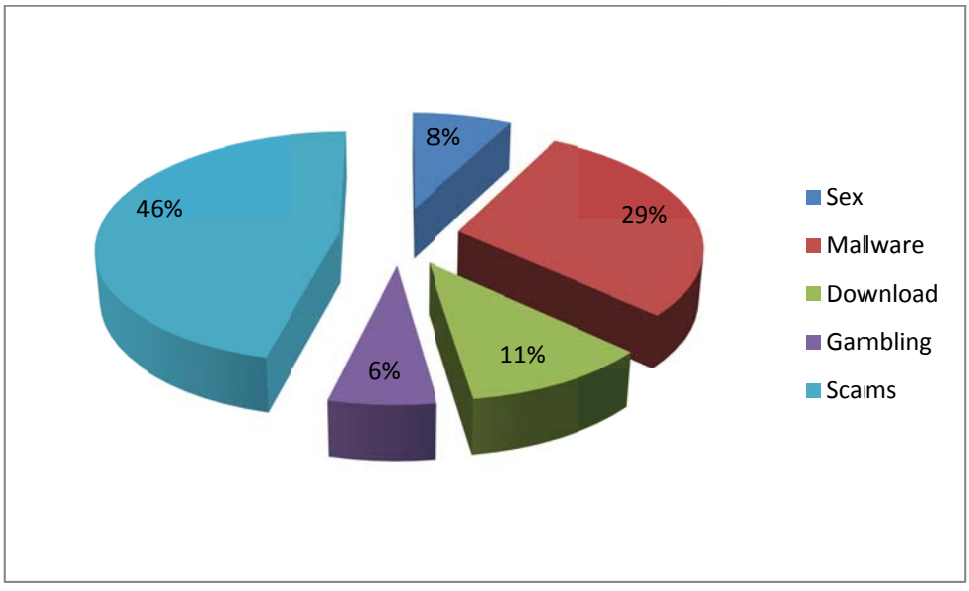
High-Risk Advertising – (Hollywood Movies/TV)

Table 3 shows the breakdown of the most common ad categories for High Risk ads across all networks. Each advertisement was downloaded, visually inspected and categorised. The results indicate that the sex industry, malware, downloading sites, gambling or scams (including employment, investment and SMS premium rate) were the most popular distinct advertising types. The categories are summarised in Figure 1.

Table 3. Frequency by Ad Category – High Risk Ads

	Sex	**Malware**	**Download**	**Gambling**	**Scams**
N	126	1243	411	32	69
%	6.70%	66.08%	21.85%	1.70%	3.67%

Figure 1. High-Risk advertising



Mainstream Advertising- All Sites (Hollywood Movies/TV)

Table 4 contains the results of the step 8 results obtained by visually inspecting every advertisement in the sample (comprising 10 pages from each of the Google Ad Transparency Top 500 complaints) to identify whether it contained any Mainstream advertising. Typically, a rogue site will have 3-4 ad panels, and in many cases, the ads were tailored to the local geographic context. In some cases, advertisements were blocked with an image stating the site was "blocked" for offshore users, indicating further evidence of geographic customisation for the advertising content. In some cases, domains associated with file sharing were "parked" and advertising displayed, even if no infringing content was actually displayed – especially where such sites had terms like "warez", "anon" and "rapidshare" in their domain name.

Only 8.24% of the ads sampled consistently showed evidence of targeting Malaysian users through the presentation of Mainstream advertising. Some ads and/or advertisers were only detected once. In a sense, this represents a type of leakage, since the Mainstream ads were a minority of the overall ads displayed (which were overwhelmingly High-Risk). A breakdown by industry category is shown in Figure 2, and the relative composition of Mainstream to High-Risk ads is shown in Figure 3.

Table 4– Mainstream Advertisers Detected

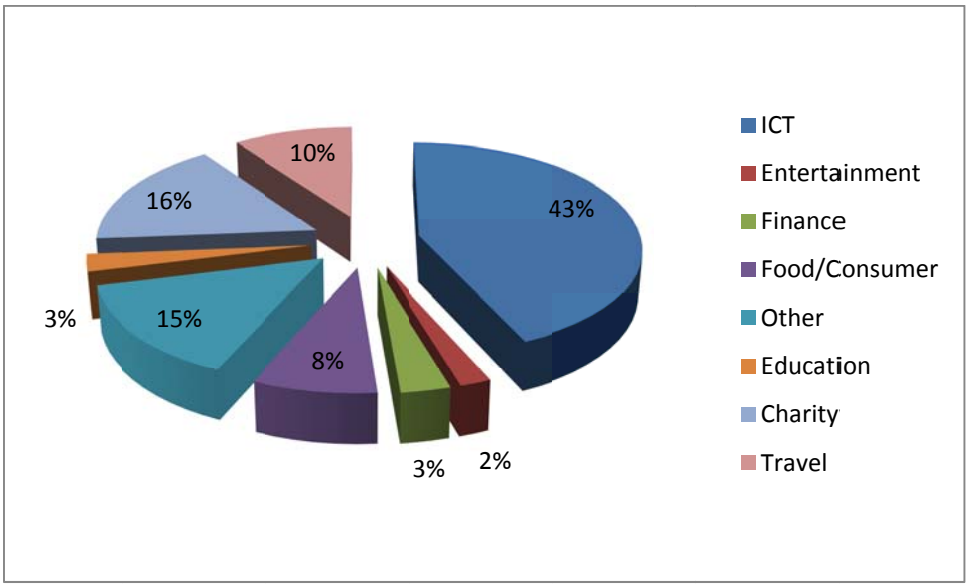


Figure 2. Mainstream advertising

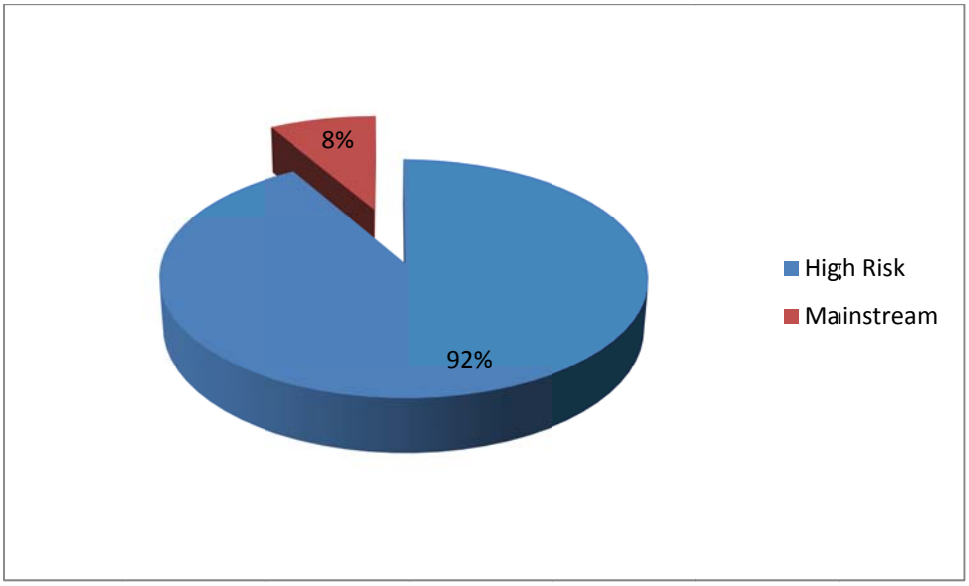


Figure 3. Mainstream versus High-Risk Advertising

Mainstream Advertising – Top 10 Ads (Hollywood Movies/TV)

Table 5 shows the frequency distribution for the ten most frequently detected Mainstream ads. The key difference to note is the relative decline in Google ads for this sample and geographic location compared to the original Australian sample, where 87% of the Mainstream ads were served up by Google ads (Watters, 2013). This may be due to local advertising conditions, network restrictions (eg, blocking of certain ad networks) or a reduction in placement of Google ads onto rogue sites as a matter of policy. Appendix C contains a list of all Mainstream advertisers for Hollywood sites.

Table 5– Mainstream Advertisers Detected (Top 10)

Advertiser	**Frequency**	**%**
Elevenia	19	11.24%
Resort World Sentosa	11	6.51%
Whitesmoke	10	5.92%
Adobe	9	5.33%
Berniaga	4	2.37%
Givology	6	3.55%
Unicef	5	2.96%
Bblibli	5	2.96%

Zalora	4	2.37%
Forestry US	4	2.37%

As some ad networks increase or decrease their presence on rogue sites, other ad networks often move in to fill the void, resulting in a type of displacement. Criminological theory suggests that displacement does not necessarily always result in negative outcomes. For example, if a more serious crime type is displaced by a less harmful type, then displacement can be positive (Felson & Clarke, 1998).

Results – Malaysian Movies/TV

Appendix D contains a list of the expert-identified rogue sites which make available torrents or links to infringing content for download. From the 360 pages analysed in Step 4, representing ten page impressions from each site, a total of 2,428 advertising items and 876 visible ads were identified in Step 6[15]. 18 unique domains for advertising networks were identified, indicating an average 48.66 visible ads per network in the sample. Appendix E contains the complete list of advertising networks detected.

The analysis is presented by reviewing the High Risk ads first, followed by the Mainstream ads.

Table 6. Frequency Analysis by Advertising Network – Top 10 [16]

Advertising Network	**Frequency**	**%**
bidvertiser.com	940	41.85%
innity.net	520	23.15%
popads.net	236	10.51%
href.asia	180	8.01%
admxr.com	100	4.45%
gloadmarket.com	80	3.56%
khalstore.com	60	2.67%

[15] Advertising items include any scripts, images, spacers etc being referenced from an Adblock domain, in addition to visible ads

[16] Note that some ad networks like isohunt.com and sumotorrent.com do not display their ads outside their own domain; they are ranked highly because of the high number of DMCA complaints against their site.

propellerads.com	50	2.23%
connextra.com	40	1.78%
popcash.net	40	1.78%

High-Risk Advertising – Top 10 (Malaysian Movies/TV)

Table 7 contains a summary of the results from the Top 10 ad networks. There were 2,246 advertising items in this sub-sample. Each of these advertisements was downloaded, visually inspected and categorised. The results indicate that malware, scams (including employment, investment and SMS premium rate), and the sex industry were the most popular distinct advertising types in Malaysia for the Top 10 networks.

Table 7. High-Risk ad type frequencies by network

Ad Network	**Ad URLs**	**Distinct Ad URLs**	**Sex Industry**	**Malware**	**Downloading**	**Gambling**	**Scams**	**TOTAL**
bidvertiser.com	940	4						0
innity.net	520	1						0
popads.net	236	4						0
href.asia	180	9					2	2
admxr.com	100	8		1				1
gloadmarket.com	80	2						0
khalstore.com	60	2						0
propellerads.com	50	4		3				3
connextra.com	40	2				1		1
popcash.net	40	1						0
TOTAL	2246	37	0	4	0	1	2	7

High-Risk Advertising – All Sites (Malaysian Movies/TV)

Table 8 shows the breakdown of the most common ad categories for High Risk ads across all networks. Each advertisement was downloaded, visually inspected and categorised. The results indicate that the sex industry, malware, gambling or scams (including employment, investment and SMS premium rate) were the most popular distinct advertising types. The categories are summarised in Figure 1. High-Risk ads accounted for 23.94% of all visible ads.

Table 8. Frequency by Ad Category – High Risk Ads

	Sex	Malware	Download	Gambling	Scams
N	10	38	14	8	60
%	7.69%	29.23%	10.77%	6.15%	46.15%

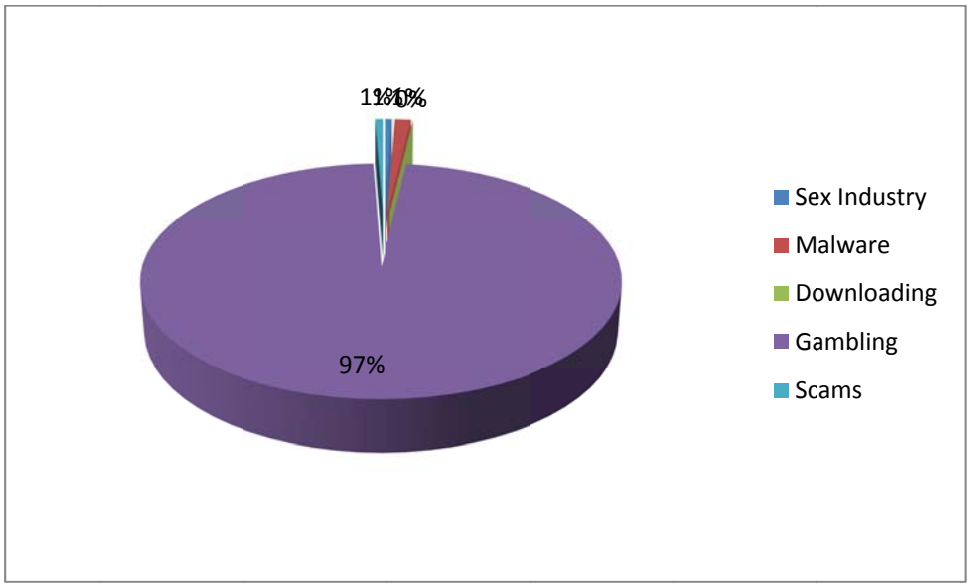


Figure 4. High-Risk advertising

Mainstream Advertising- All Sites (Malaysian Movies/TV)

Table 9 contains the results of the step 8 results obtained by visually inspecting every advertisement in the sample (comprising 10 page impressions from each of the rogue sites) to identify whether it contained any Mainstream advertising. Some 76.06% of the ads sampled consistently showed evidence of targeting Malaysian users through the presentation of Mainstream advertising. A breakdown by industry category is shown in Figure 5, and the relative composition of Mainstream to High-Risk ads is shown in Figure 6.

Table 9– Mainstream Advertisers Detected

Advertiser	**Frequency**
US Forest Service	84
Geng Tube	38
The Economist	26
adf.ly	25
Lice Ninja Bangi	21
International University and Colleges	21
Proride	20
PopAds	20
Moviehouse	14
DV Lottery Services	13
Projek IQRA	10
L'Oreal	10
Mydeal Malaysia	10
Jobdirumah.com	10
Logo Books	10
Perodua	9
Maybank	8
Discover the Forest	7

Singapore Airlines	6
Nissan	6
FEMA	6
US Red Cross	5
Givology	5
Bio Velocity	5
Maybank	3
Teach.org	3
Planet49	2
Sentosa Resorts World	2
Fat Fx	2
Daily Fitness Centre	2
United Way	1
National Institute of Health	1
Fatherhood.gov	1
Feeding America	1
NHTSA	1
Bedsider.org	1
Rushia Bun Pau & Pizza	1
Yourvoucher Malaysia	1
Adoptuskids.org	1
Television Fanatic	1

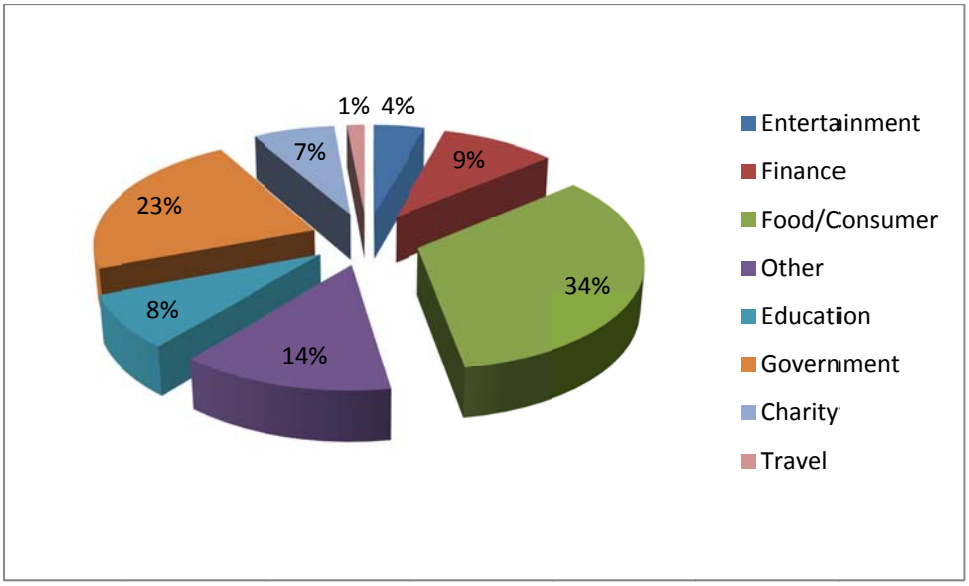


Figure 5. Mainstream advertising

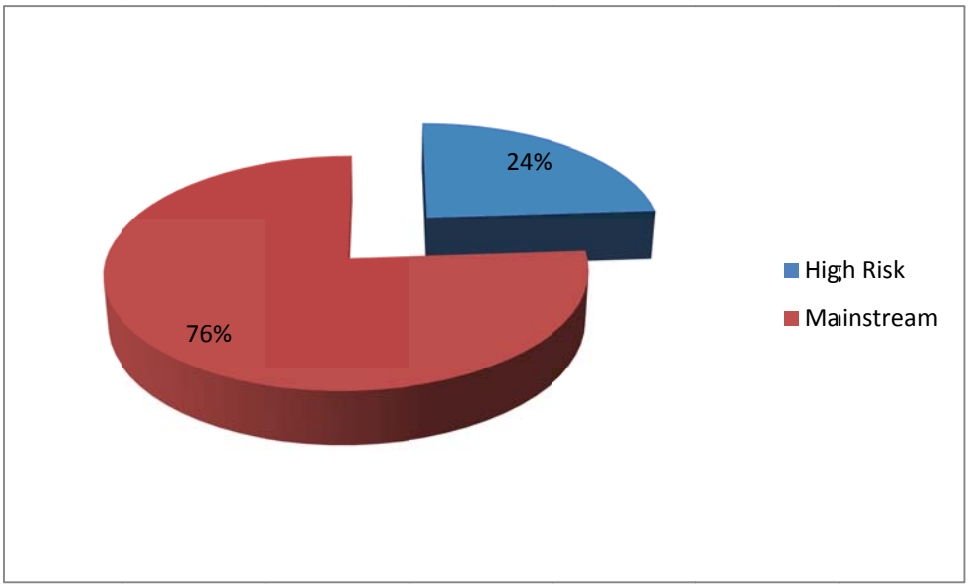


Figure 6. Mainstream versus High-Risk Advertising

Mainstream Advertising – Top 10 Ads (Malaysian Movies/TV)

Table 10 shows the frequency distribution for the ten most frequently detected Mainstream ads.

Table 10 – Mainstream Advertisers Detected (Top 10)

Advertiser	**Frequency**	**%**
US Forest Service	84	20.34%

Geng Tube	38	9.20%
The Economist	26	6.30%
adf.ly	25	6.05%
Lice Ninja Bangi	21	5.08%
International University and Colleges	21	5.08%
Proride	20	4.84%
PopAds	20	4.84%
Moviehouse	14	3.39%
L'Oreal	10	3.15%

Conclusion

The hypothesis that local content would attract more mainstream advertising than Hollywood content was supported by this study, as was the case for Hong Kong. While the Taiwan study found that this was true for local Chinese content, this study provides the second direct comparison of advertising for local content versus Hollywood, and the results are startling: while the levels of mainstream advertising on rogue sites promoting Hollywood content was comparable with other countries including Australia, New Zealand, Canada, and Singapore, the levels of mainstream increased by several times when local content was promoted. This may be a result of brand protection strategies being more focused on sites which are listed in Google's DMCA Transparency Report; in a sense, the local content sites in Malaysia appear to be flying "under the radar". Also, while some local content sites were carrying mainstream ads by local networks, Yahoo! and Google ads were also detected promoting mainstream content.

The key findings from the analysis of this first-ever Malaysian data set are discussed below:

- For local content, 72.1% of movie and TV ads were Mainstream, while 27.9% were High Risk. In contrast, 8.24% of Hollywood ads were Mainstream, while 91.76% were High Risk.
- For Hollywood content, high-risk advertising primarily comprised malware, while for local content, the greatest risk was from scams.
- Advertising for the sex industry was higher for local sites than Hollywood sites.
- The top ad networks serving ads to Malaysian residents for local content included bidvertiser.com, innity.net and popads.net, while adcash.com, adexprt.com and depositfiles.com were found for Hollywood sites.
- A number of household name brands in Malaysia choosing to advertise on sites and their pages which are promoting the distribution of local and Hollywood infringing content

(movies and TV shows). Further investigation is needed to uncover the mechanics of how these ads are selected to appear; are advertisers engaging directly with ad networks, or are ad networks operating at a wholesale level and distributing ads to other networks through a resale programme? Who, eventually, has control over the display of this type of advertising space?

- Some advertisers may be unwittingly placing ads which contain no obvious textual references to piracy, yet these pages do contain links to infringing content, and their referring pages have been verified by Google as DMCA-infringing.

Drawing together these findings, some key lessons can be drawn:

- Malaysian residents have a higher chance of viewing Mainstream ads on local movie/TV sites compared to Hollywood sites; the converse is also true.
- Malaysian ads do not appear to be filtered, as sex industry ads were quite prevalent. Malaysia should investigate applying further controls that are text based as well as image based (eg, Ho & Watters, 2004).
- Advertisers need to have better mechanisms to control where their ads are eventually displayed on ad networks. Better systems for operational assurance and detection of misplaced ads need to considered, whether they operate using a whitelist or a blacklist (Ho & Watters, 2005). Mechanisms should be created to build awareness for advertisers to be aware of where their ads are being placed, and they should be given an opportunity to respond ethically.
- Regulatory approaches need to be considered to control the revenue flowing to rogue websites, and to minimise harm to users, if raising awareness amongst advertisers proves ineffective. A proposed code of conduct (Dredge, 2013) would be a first step to isolating rogue websites. Advertisers recently succeeded in pressuring Facebook, for example, to remove offensive ad placements by threatening to remove ads as a group (Cellan-Jones, 2013).Advertisers may be resistant to further regulation, since this will reduce revenue.
- Regulators in Malaysia should investigate the online advertising industry, to identify and distinguish between the genuine mainstream ad industry and the non-regulated players, in order to block the unwanted online advertisements that are freely advertising illegal products/services in Malaysia.
- Other types of rogue content have been managed effectively by legal sanctions in the past. For example, paid search results for pharmaceuticals without prescriptions (O'Donnell, 2013) were removed by Google after they paid a very significant fine. However, Google's organic search results continue to display results from rogue drug sellers, ranging from marijuana through to MDMA and ecstasy (Watters & Phair, 2012). Searching Google for "buy ecstasy" returns numerous pages such as http://buyecstasypillsonline.wordpress.com/2013/07/27/buy-cheap-ecstasy-pills-online/ where users can order illicit drugs and have them delivered to order. Regulation of this type of advertising can be effective but more needs to be done.
- Since cyber criminals are very effective at exploiting jurisdictional differences, a global, industry wide code may have a greater impact on revenue flows for rogue websites. However, industry codes need to engage with ad networks that are placing ads for High Risk advertisers. At this stage, none of the top advertising networks supporting rogue websites

appear to be involved in the proposed code of conduct[17]. Also, no additional burden should be placed on rightsholders to police the internet for offensive material.

- Finally, and perhaps most importantly, parents and educators need to be aware of that the sex industry and online gambling sites specifically target torrent search and file locker sites for advertising their services. Ads promoting gambling compromised were particularly prevalent on sites promoting local content. There are absolutely no age warnings on these pages, and no attempt is made by the Pirate Bay (for example) to verify if users are adults. Parents need to be aware that this is the type of content that will be served up to their children, even if they are only intending to download torrent for music or less offensive content. The absence of traditional regulatory mechanisms for effectively controlling online content mean that new subcultural norms are rapidly being established online, and these can have profoundly negative consequences; for example, a progression model of rising interest in child exploitation material has been linked to the rise of the online porn culture, particularly where young users are inadvertently exposed to pornography through advertising (Prichard et al, 2013).

References

Browser Media (2008). DoubleClick deal means Google controls 69% of the online ad market. Downloaded from http://www.browsermedia.co.uk/2008/04/01/doubleclick-deal-means-google-controls-69-of-the-online-ad-market/

Cameron, N. (2013). Hong Kong's online advertising market valued at $17.1bn. Downloaded from http://www.cmo.com.au/article/466022/Hong Kong_online_advertising_market_valued_17_1bn/

Cellan-Jones, R. (2013). Facebook removes ads from controversial pages to avoid boycott. Downloaded from http://www.bbc.co.uk/news/technology-23097411

Cleland, S. (2013). Why Google is Big Brother Inc. – A One-Page Graphic. Downloaded from http://www.precursorblog.com/?q=content/why-google-big-brother-inc-%E2%80%93-a-one-page-graphic-part-33-google-disrespect-privacy-series

Detica (2012). A data driven study of websites considered to be infringing copyright. Downloaded from http://www.prsformusic.com/aboutus/policyandresearch/researchandeconomics/Documents/TheSixBusinessModelsofCopyrightInfringement.pdf

Dredge, S. (2013). Google, David Lowery and the BPI talk ad-funded piracy. Downloaded from http://musically.com/2013/05/28/live-google-david-lowery-and-the-bpi-talk-ad-funded-piracy/

eMarketer (2012). US Online Advertising Spending to Surpass Print in 2012. Downloaded from http://www.emarketer.com/Article/US-Online-Advertising-Spending-Surpass-Print-2012/1008783

Felson, M and Clarke RV (1998). Opportunity Makes the Thief: practical theory for crime prevention Police Research Series Paper 98. London: Home Office

[17] http://www.bbc.co.uk/news/technology-23325627

Gopal, R. D., Sanders, G. L., Bhattacharjee, S., Agrawal, M., & Wagner, S. C. (2004). A behavioral model of digital music piracy. Journal of Organizational Computing and Electronic Commerce, 14(2), 89-105.

Herps, A., Watters, P.A. & Pineda-Villavicencio, G. (2014). Measuring the prevalence of behavioral advertising using tracking cookies. *Proceedings of the 4th Cybercrime and Trustworthy Computing Workshop.*

Ho, W. H., & Watters, P. A. (2004, October). Statistical and structural approaches to filtering internet pornography. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on* (Vol. 5, pp. 4792-4798). IEEE.

Ho, W. H., & Watters, P. A. (2005, April). Identifying and blocking pornographic content. In *Data Engineering Workshops, 2005. 21st International Conference on* (pp. 1181-1181). IEEE.

Kuprijanko, A. (2009). Försvaret: verksamheten är laglig. Sydsvenskan. Downloaded from http://archive.is/omksR.

McStay, Andrew. The mood of information: a critique of online behavioural advertising. Continuum, 2011.

Medoff, Norman J. Just a click away: Advertising on the Internet. Allyn & Bacon, Inc., 2000.

Olsson, S. (2006). Pirate Bay drar in miljonbelopp. Svenska Dagbladet. Downloaded from http://www.svd.se/nyheter/inrikes/pirate-bay-drar-in-miljonbelopp_334410.svd

Prichard, J., Spiranovic, C., Watters, P.A. & Lueg, C. (2013). Young people, child pornography, and subcultural norms on the Internet. Journal of the American Society for Information Science and Technology, 64, 992-1000.

Rawls, J. (1999). A Theory of Justice. Belknap Press.

Sheehan, B., Tsao, J., Bruno, E., Crider, D., Cutrone, J., Jones, C., & Serra, A. (Submitted). Improving the effectiveness of anti-digital music piracy advertising to college students.

Stabek, A., Brown, S., & Watters, P. A. (2009, July). The Case for a Consistent Cyberscam Classification Framework (CCCF). In Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC'09. Symposia and Workshops on (pp. 525-530). IEEE.

Stabek, A., Watters, P., & Layton, R. (2010, July). The seven scam types: mapping the terrain of cybercrime. In Cybercrime and Trustworthy Computing Workshop (CTC), 2010 Second (pp. 41-51). IEEE.

Sundberg, S. (2009). TPB har tjänat tio miljoner om året" (blog) (in Swedish). Svenska Dagbladet. Downloaded from http://www.webcitation.org/6D8mmNnUX

Taplin, J. (2013). USC Annenberg Lab Ad Transparency Report – January. Downloaded from http://www.annenberglab.com/sites/default/files/uploads/USCAnnenbergLab_AdReport_Jan 2013.pdf

Treverton, G., Matthies, C., Cunningham, K., Goulka, J., Ridgeway, G., & Wong, A. (2009). Film Piracy, Organized Crime and Terrorism. RAND Corporation. Downloaded from http://www.rand.org/content/dam/rand/pubs/monographs/2009/RAND_MG742.pdf

Watters, P.A. (2014a). A Systematic Approach to Measuring Advertising Transparency Online: An Australian Case Study. In S. Cranefield, A. Trotman, & J. Yang (Eds.)Vol. 155 (pp. 59 - 67). *Proceedings of the Second Australasian Web Conference*

Watters, P.A. (2014b). Sweet as? Advertising on rogue websites in New Zealand. http://papers.ssrn.com/abstract=2466696

Watters, P.A. (2014c). Mainstream Ad Support for Online Piracy in Taiwan. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2405281

Watters, P.A. (2014d). Mainstream Advertising on Rogue Websites in Hong Kong: A Comparison of Chinese and Western Titles. http://papers.ssrn.com/abstract=2468700

Watters, P.A. (2013a). The Prevalence of High-Risk and Mainstream Advertisements Targeting Canadians on Rogue Websites (December 2, 2013). Available at SSRN: http://ssrn.com/abstract=2389850

Watters, P.A. (2013b). Measuring Advertising Transparency in Singapore: An Investigation of Threats to Users. Available at SSRN: http://ssrn.com/abstract=2362626 or http://dx.doi.org/10.2139/ssrn.2362626

Watters, P.A. (2012). Taming the Cookie Monster: How Companies Track us Online. Centre for Internet Safety, University of Canberra. ISBN 978-1-922017-04-8.

Watters, P.A. & Phair, N. (2012). Detecting Illicit Drugs on Social Media Using Automated Social Media Intelligence Analysis (ASMIA). CSS 2012: 66-76.

Womack, B. (2013). Google Is Projected to Expand Lead in Online-Ad Market. Downloaded from http://www.bloomberg.com/news/2013-06-13/google-is-projected-to-expand-lead-in-online-ad-market.html

Appendix A – Top 500 DMCA Notices for Movies and TV

TO BE INSERTED

Appendix B – Full List of advertising networks detected (Hollywood Movies and TV)

Frequency	Ad Network
propellerads.com	802
filestube.to	402
adcash.com	213
disqus.com	122
xtendmedia.com	120
depositfiles.com	99
adexprt.com	97
uasdel.com	92
onclickads.net	92
twitter.com	86
ismcorp.vn	84
popads.net	80
facebook.com	70
torrentroom.com	56
fhserve.com	56
beta.filestube.to	56
totaladperformance.com	54
propellerpops.com	48
movie4k.to	42
admxr.com	41
adserve.com	40
wikipedia.org	40
mads2srv.info	39
yllix.com	35
cdn-image.com	32
adsopx.com	32
liveadoptimizer.com	31
ladssrv.info	29
tlvmedia.com	28
torrage.com	28
zoink.it	28
rtbpop.com	27
w3.org	26
sumotorrent.sx	25
ads4vn.com	24
netsprint.eu	22
googletagservices.com	21
adshost2.com	21
thepiratebay.org	21
adbooth.net	20
cpxinteractive.com	19

tv-release.net	19
freefilesdownloader.com	19
torrentday.com	18
cpmterra.com	18
ad-center.com	18
torrenthound.com	18
yourbittorrent.com	18
argentinawarez.com	17
velmedia.net	16
acquirethisname.com	15
tvrage.com	15
zap2it.com	14
thefutoncritic.com	14
imdb.com	14
engine.4dsply.com	13
google-analytics.com	13
adshost1.com	12
torrentdownloads.net	12
torrentreactor.net	12
aclantis.com	12
google.com	11
monova.org	11
adtransfer.net	11
kickasstorrents.com	11
zwaar.org	11
arbopl.bbelements.com	10
usanetwork.com	10
awempire.com	10
btjunkie.org	10
extra33.com	10
newtorrents.info	9
fulldls.com	9
downloaddegraca.com	9
torrentzap.com	9
bitsnoop.com	8
lzjl.com	8
yesads.com	8
yieldmanager.com	8
torrentfunk.com	8
directdownload.tv	7
wuper.net	7
fileserve-movies.com	7
alivetorrents.com	7
fenopy.com	7

github.com	6
darkmachine.pl	6
rarbg.com	6
BitTorrent.AM	6
extratorrent.com	6
sceper.eu	6
releaseserve.info	6
mediafire.com	6
puraseries.org	6
linksfu.com	6
rlsbb.com	6
linktrooper.com	6
nuseek.com	6

Appendix C – Full List of Mainstream advertisers on Hollywood sites

Advertiser	**N**	**%**
Elevenia	19	11.24%
Resort World Sentosa	11	6.51%
Whitesmoke	10	5.92%
Adobe	9	5.33%
Berniaga	8	4.73%
Givology	6	3.55%
Unicef	5	2.96%
Bblibli	5	2.96%
Zalora	4	2.37%
Forestry US	4	2.37%
AOL	4	2.37%
Discover the Forest	3	1.78%
SingTel	3	1.78%
The Economist	3	1.78%
Redcross US	3	1.78%
Mastercard	3	1.78%
Techsmith	2	1.18%
Closeup	2	1.18%
Pizza Hut	2	1.18%
Apple	2	1.18%
Travelok	2	1.18%
Etsy	2	1.18%
Lancia	2	1.18%
Emirates	2	1.18%
Conbaoso	2	1.18%
Adoptuskids	2	1.18%
Avira	2	1.18%
CompTia	1	0.59%
Imagine Publishing	1	0.59%
GloCal	1	0.59%
Stockphoto	1	0.59%
Wordpress	1	0.59%
TPBank	1	0.59%
Wild Web Lab	1	0.59%
Softlayer	1	0.59%
National University of Singapore	1	0.59%
Lazada	1	0.59%
lp3i	1	0.59%
datviet.mobi	1	0.59%
Visual	1	0.59%

Clinton Foundation	1	0.59%
Beck Sapphire	1	0.59%
American Express	1	0.59%
Ashampoo	1	0.59%
Windows	1	0.59%
Tokobagus	1	0.59%
EF Academy	1	0.59%
Fossil	1	0.59%
Reebonz	1	0.59%
Easyuni	1	0.59%
Senselock	1	0.59%
Canon	1	0.59%
Cleanmymac2	1	0.59%
YAC	1	0.59%
Airbnb	1	0.59%
Shell	1	0.59%
ASUS	1	0.59%
AV1solutions	1	0.59%
US Heart Foundation	1	0.59%
Interia	1	0.59%
Modomania	1	0.59%

Appendix D – Rogue Sites for Malaysian Movies/TV

TO BE INSERTED

Appendix E – Full List of advertising networks detected (Malaysian Movies/TV)

Advertising Network	**Frequency**
bidvertiser.com	940
innity.net	520
popads.net	236
href.asia	180
admxr.com	100
gloadmarket.com	80
khalstore.com	60
propellerads.com	50
connextra.com	40
popcash.net	40
iasrv.com	30
agenaffiliate.com	30
adk2.com	30
smowtion.com	22
onclickads.net	20
1stdrama.com	20
downloadur.org	20
adcash.com	10