CONTENT	LEXICAL	LINK	NETWORK	DNS Response	REACTIVE	OTHERS	JAVASCRIPT
iframe	Url length	Maximum domain link ratio	redirect count	Resolved IP count	**Exploit kit used**	entropy	eval()
script tag(for exploit kits; features of common script codes)	number of special characters(RATIO)	phishing link ratio	download packet content length	NameServer count	**Injected content**	**ActiveX objects**	setTimeOut()
redirection check	path length	spamming link ratio	actual download bytes	malicious ASN	**Obfuscation of script**	**ActiveX executions**	setInterval()
line count	number of digits	**malware link ratio [1]**	domain lookup time	WhoIS Registration		3-4-5 n grams	**in-built functions for deobfuscation**
hyperlink count	number of alphabets	link popularity	download speed	IP Address Location		TF-IDF	**pieces of code having deobfuscation routine**
script content in page(%)	number of keys in query	Time for retrieving content [2	details of CA issurer	Whether source and destination are same		**spam assasin plugin**	entropy of strings
script function count	inclusion of http(0 or 1)		start/end date of Certificate	DNS Response			entropy of script
web page size	presence of IP IN URL		Presence of Top Level Domain(TLD	TTL(is shortlived)			long strings
HTML tag count	domain token count		Presence of Sub-domain	Geolocation of IP [3]			max entropy of string
Applet tag	longest domain token count		**Download without warning**	IP of mailserver, reverse IP			long variables/functions
embed tag	longest path token count		**Redirection without warning**	nameserver IP			string direct assignment
XML tag	average count		host component count in page	resolved PTR record			**string modification function**
style tag	brand presence [4]		TCP packets	if PTR record = A record [5]			event attachment
for tag	length of url		distinct TCP ports	for each of A,MX,NS records ->first IP returned is same/not			fingerprinting functions
meta tag	number of dots		**remote IPs(not including DNS ser**	target of redirections			suspicious objects
img tag	whether blacklisted word is present		**urgent TCP packets**				suspicious strings
src tag	hyphens used instead of dots in domain name		UDP packets				**DOM modification functions**
header tag (Values of all the variables in header as attributes)	another char used instead of dots in hostname		Average packet rate				average length of strings
max, min and median size of every file extension	length of directory						whitespace % of script
character count	number of sub-directory tokens						length of strings
script with wrong extension	max number of dots and others in sub-directory token						average script line length
number of elements with smaller area	length of argument						strings containing 'iframe' tag
div tag	number of variables						strings containing suspicious tags
length of page name	length of the longest variables						length of script in chars
unknown tags	if domain part exists in Alexa rank						ratio of string definition and uses
elements containing suspicious content	subdomain length						length of string passed to eval
suspicious objects	@’ and ’-’ count						**instantiated components**
% whitespace	Kolmogorov-Smirnov statistic						sequence of method calls
meta refresh tag	Kullback-Leibler divergence						**Number of user prompts**
object tag	Suspicious words count						token of user prompts
elements with source in external domain	**Euclidean distance**						**pop-up windows**
out of place elements	length of filename in URL						behaviour of pop-up window
presence of double documents	presence of port number in URL						**plugins**
number of same origin links	relative length of URL						application type of plugins
number of different origin links	number of special characters						**MIME injection [6]**
number of external Javascript files							suspicious unicode characters
changed DOM's location							**suspicious decoding results**
delimiters							overlong decoding results
							dangerous element creation
							URI/CLSID in attribute setter [7]
			Rating	**Colour**			**dangerous tag injection via innerHTML setting**
			Phase 1	🟨			**the number of obfuscation functions**
			Phase 2	🟦			**event triggering functions**
			Phase 3	⬜			

[1] the ratio from domains of a specific malicious type

[2] Pagespeed API google

[3] country code, region, time zone, netspeed

[4] Brand names can be taken from the SLDs of the Alexa [2] top 500 site list.

[5] For a particular IP

[6] https://pdfs.semanticscholar.org/89d1/633f0019ff2d561132a29fa5a9ab549fa8bd.pdf

A script applies a MIME type that is potentially dangerous to an existing DOM object such as application/java-deployment-toolkit.

[7] Iceshield.pdf