IPViking Data Guide

1/22/14

CONTENTS

1. Summary 5
2. IPViking API Data Elements 6
2.1. Examples 6
IPViking IP Search Tool 6
IPViking API 7
2.2. IPQ Score 11
2.3. Factoring 11
IPViking Category Factor 11
Data Age Factor 12
IP Resolve Factor 12
ASN Record Factor 12
ASN Threat Factor 12
BGP Delegation Factor 13
IANA Allocation Factor 13
Country Risk Factor 13
Region Risk Factor 13
Geomatch Factor 14
Geomatch Distance 14
IPViking Geofilter Factor 14
IPViking Geofilter Rule 15
Search Volume Factor 15
IPViking Personal Factor 15
2.4. Context / Rationale (Categories-Protocols) 15
Anonymous Proxies 16
Bogons 16
Bots 17
Malware URL 18
Passive DNS 18
HTTP 19
Timestamps 19
2.5. Network Telemetry 20

IP Address and Hosts 20
Service Providers/Autonomous System/Organization 20
2.6. Geographic Intelligence 20
2.7. Device Type 20
3. Applying IPViking Data 21
3.1. Use Cases 21
Fraud 21
Payments 21
Other 22
Application 22
Login 22
Other 22
In-line Network Security (Prevention) 23
SOC / NOC (Detection I) 23
Forensics and Incident Handling (Detection II) 23
Analytics (Cyber Teams) 23
3.2. Protocol 24
3.3. Point in the Network 24
4. Operations 25
4.1. Overview 25
4.2. Considerations 25
4.2.1 Internal Devices 25
4.2.2 Communication Direction 25
4.2.3 Communication Protocol 25
4.2.4 Norse Intelligence on External Device 26
4.2.5 Expected Behavior 26
4.2.6 Exceptions 26
4.2.7 Action 26
4.3. The Living Document 27
4.4. How Norse Can Help 27
4.5. Example Rules 27
5. Capabilities 28

Appendix A: Categories and Protocols 29
Categories 29
Protocols 31

All information provided by Norse Corporation, in any form, is proprietary to Norse Corporation and is protected by U.S. and foreign laws governing intellectual property. All such information published by Norse Corporation or presented by its employees is copyright-protected, inclusive of material appearing in a hard-copy format, electronically, or on our Web site. Because law does not require a copyright notice, the omission of the copyright notice by Norse Corporation does not invalidate copyright protection and does not indicate that Norse Corporation authorizes the reproduction of such proprietary material.

Acceptable Use
Acceptable Use is a privilege that allows users to make copies of copyrighted information without the specific consent of Norse Corporation in specific limited situations.

Internal presentations
Norse Corporation considers Acceptable Use to be the reproduction of portions of Norse Corporation Information's research for use in a presentation to individual employees within the same company. If the research involves the use of graphs/charts, we request that the content be reproduced exactly as it appears with no modifications and that the source line includes the information supplied in our graphics/charts as well as the name Norse Corporation Information, if Norse Corporation Information is not already identified as the source. Employees who receive any of the limited number of copies of these presentations should be informed that the copies are not for external use outside the company and not to be provided to non-employees, independent contractors, or third-party consultants.

What is Not Considered Acceptable Use
Uses that conflict with Norse Corporation Information's business practices or impair the market for Norse Corporation Information's materials are not considered Acceptable Use. Unacceptable uses include, without limitation:

- Copying of complete reports for archives, files, or otherwise, such as to avoid purchasing a reprint.
- Posting complete or partial sections of documents on an Internet site or on an intranet site without purchasing a separate Intranet Copy of the report (see Multi-User Files and Intranet Copies, above).
- Scanning or otherwise importing publications into an electronic storage/retrieval system.
- Distribution of single-user copies of reports to anyone other than the single user, including employees of the organization, either on hard copy or through electronic data transmission systems such as electronic mail.
- Distribution of multi-user copies of reports beyond the number and scope outlined in the agreement sent with the multi-user copy.
- Distribution of publications to anyone outside the organization through any medium or format, including but not limited to hard copy or electronically such as via electronic mail.
- Distribution to company salespeople who could use the publications for discussions with customers or prospective customers.
- Using research or excerpts from research in public relations or sales campaigns.

APPROVAL MAY BE REQUIRED
If a desired use of Norse Corporation Information's material is not covered under Acceptable Use, then a request for Norse Corporation Information's approval is required. Requests to quote, excerpt, reproduce, or redistribute Norse Corporation Information's materials should be submitted to Sam Glines, Chief Executive Officer. sg@norse-corp.com or 314-480-6447.

1. SUMMARY

This document provides understanding of Norse dark intelligence and how it can be applied to your environment based on scenarios. It is meant to be a guide for security operations personnel and engineers, as well as evaluators of Norse API or data products.

2. IPVIKING API DATA ELEMENTS

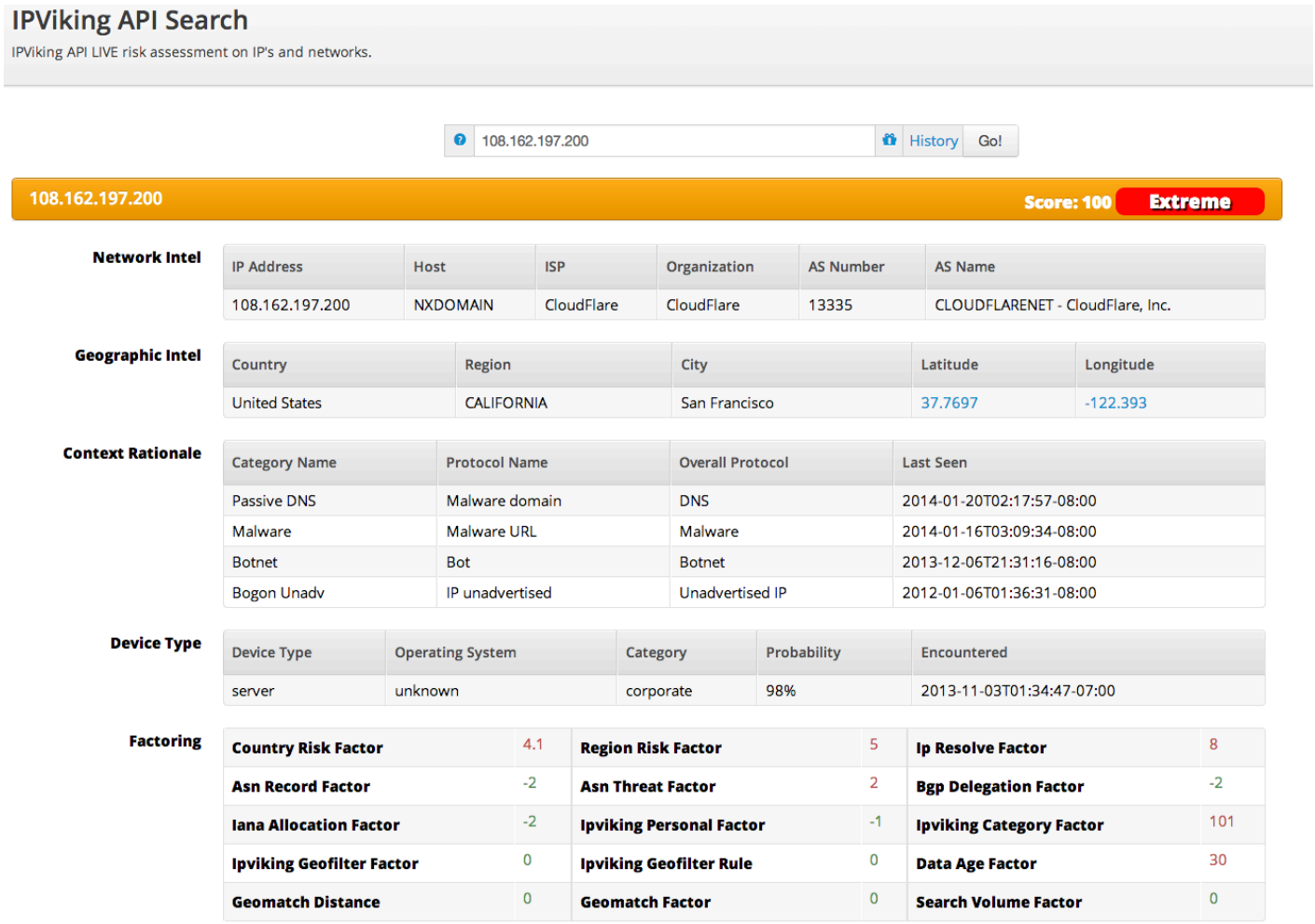
This section provides field-level descriptions of IPViking API data elements, which are a superset of the Darklist API dark elements.

2.1. Examples

IPViking API is a REST API that returns results in JSON or XML format. The IPViking API is also called by the ipviking.com portal IP Search Tool and displays the same data elements. Both JSON output and screenshots will be used to throughout this document.

Here is output for IP address, 108.162.197.200. Note scores may be different should you query this same IP as age of activity is a key factor in score calculation.

IPVIKING IP SEARCH TOOL



Network Intel

IP Address	Host	ISP	Organization	AS Number	AS Name
108.162.197.200	NXDOMAIN	CloudFlare	CloudFlare	13335	CLOUDFLARENET - CloudFlare, Inc.

Geographic Intel

Country	Region	City	Latitude	Longitude
United States	CALIFORNIA	San Francisco	37.7697	-122.393

Context Rationale

Category Name	Protocol Name	Overall Protocol	Last Seen
Passive DNS	Malware domain	DNS	2014-01-20T02:17:57-08:00
Malware	Malware URL	Malware	2014-01-16T03:09:34-08:00
Botnet	Bot	Botnet	2013-12-06T21:31:16-08:00
Bogon Unadv	IP unadvertised	Unadvertised IP	2012-01-06T01:36:31-08:00

Device Type

Device Type	Operating System	Category	Probability	Encountered
server	unknown	corporate	98%	2013-11-03T01:34:47-07:00

Factoring

Country Risk Factor	4.1	**Region Risk Factor**	5	**Ip Resolve Factor**	8
Asn Record Factor	-2	**Asn Threat Factor**	2	**Bgp Delegation Factor**	-2
Iana Allocation Factor	-2	**Ipviking Personal Factor**	-1	**Ipviking Category Factor**	101
Ipviking Geofilter Factor	0	**Ipviking Geofilter Rule**	0	**Data Age Factor**	30
Geomatch Distance	0	**Geomatch Factor**	0	**Search Volume Factor**	0

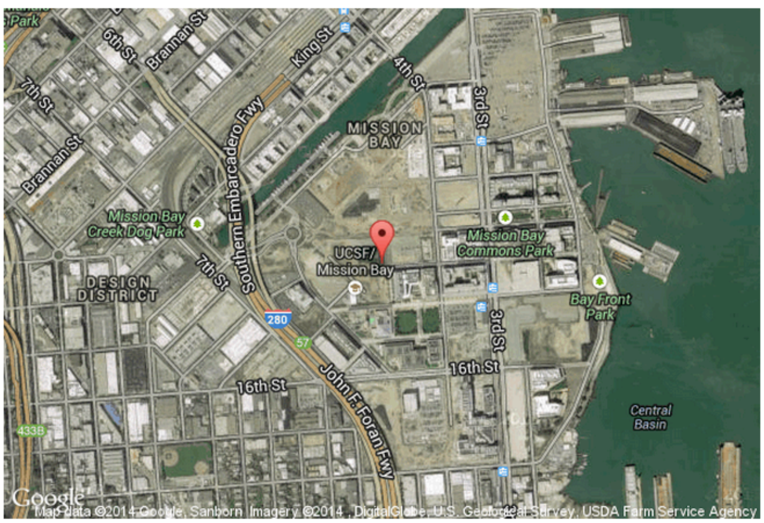
Malware URLs

Category Name	MIME type	URL	Last Seen
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v1688/?vug&a...	2013-10-24T02:36:30-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:07-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00

Passive DNS

Domain	Malware / Total URLs	Last Seen
moousi.com	0 / 1	2014-01-20T02:17:57-08:00
applicationscreditforally.asia	19757 / 21840	2013-10-24T02:14:12-07:00
www.cuzzsoft.com	0 / 1	2013-10-02T09:42:37-07:00
cuzzsoft.com	0 / 1	2013-10-02T09:42:41-07:00
greatfilesarey.asia	140 / 174	2013-10-01T23:16:03-07:00

Map



IPVIKING API

A call to the IPViking API results in the same data elements.

```
{
   "response" : {
      "risk_color" : "red",
      "entries" : [
         {
            "category_factor" : "50",
            "category_id" : "81",
            "last_active" : "2014-01-16T03:09:34-08:00",
            "protocol_id" : "392",
```

```
            "overall_protocol" : "Malware",
            "category_name" : "Malware",
            "protocol_name" : "Malware URL"
         },
         {
            "category_factor" : "20",
            "category_id" : "82",
            "last_active" : "2014-01-20T02:17:57-08:00",
            "protocol_id" : "393",
            "overall_protocol" : "DNS",
            "category_name" : "Passive DNS",
            "protocol_name" : "Malware domain"
         },
         {
            "category_factor" : "65",
            "category_id" : "5",
            "last_active" : "2013-12-06T21:31:16-08:00",
            "protocol_id" : "41",
            "overall_protocol" : "Botnet",
            "category_name" : "Botnet",
            "protocol_name" : "Bot"
         },
         {
            "category_factor" : "25",
            "category_id" : "2",
            "last_active" : "2012-01-06T01:36:31-08:00",
            "protocol_id" : "31",
            "overall_protocol" : "Unadvertised IP",
            "category_name" : "Bogon Unadv",
            "protocol_name" : "IP unadvertised"
         }
      ],
      "ip" : "108.162.197.200",
      "clientID" : 0,
      "urlview" : [
         {
            "malware" : "adware/installrex.gen to Win.Adware.Agent-2573",
            "last_seen" : "2013-10-01T23:16:03-07:00",
            "first_seen" : "2013-10-01T23:16:03-07:00",
            "url_md5" : "a047a6becb037a5d5c47e8c4c20144ea",
            "url"                                                                                    :
"http://greatfilesarey.asia/v964/?product_name=Kaplan+Schweser+CFA+2013+Level+2+%E2%80%
93+Ebooks+and+Video+Tutorials%26amp%3Binstaller_file_name%3DKaplan+Schweser+CFA+2013+Le
vel+2+%E2%80%93+Ebooks+and+Video+Tutorials%26amp%3Br%3D540",
            "mime_type" : "binary"
         },
         {
            "malware" : "adware/installrex.gen to Win.Adware.Agent-2573",
            "last_seen" : "2013-10-01T23:16:03-07:00",
            "first_seen" : "2013-10-01T23:16:03-07:00",
            "url_md5" : "8d69dc7b12c816ed388cbeabadc7749a",
            "url"        :        "http://greatfilesarey.asia/v724/?q=Anthony+Robbins+-
+The+decision+that+ensures+your+success",
```

```
      "mime_type" : "binary"
   },
…
Note: records ommitted for brevity
…
],
"timestamp" : "2014-01-20T05:55:18-08:00",
"ip_info" : {
   "autonomous_system_number" : "13335",
   "autonomous_system_name" : "CLOUDFLARENET - CloudFlare, Inc."
},
"method" : "ipview",
"geoloc" : {
   "country" : "United States",
   "region_code" : "CA",
   "country_code" : "US",
   "region" : "CALIFORNIA",
   "city" : "San Francisco",
   "longtitude" : "-122.393",
   "latitude" : "37.7697",
   "internet_service_provider" : "CloudFlare",
   "organization" : "CloudFlare"
},
"risk_name" : "Extreme",
"devicetype" : [
   {
      "probability_percent" : "98",
      "device_category" : "corporate",
      "device_type" : "server",
      "encountered" : "2013-11-03T01:34:47-07:00",
      "device_os" : "unknown"
   }
],
"dnsview" : [
   {
      "url_count" : "174",
      "domain" : "greatfilesarey.asia",
      "malware_url_count" : "140",
      "last_seen" : "2013-10-01T23:16:03-07:00",
      "first_seen" : "2013-10-01T23:16:03-07:00"
   },
   {
      "url_count" : "1",
      "domain" : "www.cuzzsoft.com",
      "malware_url_count" : "0",
      "last_seen" : "2013-10-02T09:42:37-07:00",
      "first_seen" : "2013-10-02T09:42:37-07:00"
   },
   {
      "url_count" : "1",
      "domain" : "cuzzsoft.com",
      "malware_url_count" : "0",
      "last_seen" : "2013-10-02T09:42:41-07:00",
```

```
            "first_seen" : "2013-10-02T09:42:41-07:00"
         },
         {
            "url_count" : "21840",
            "domain" : "applicationscreditforally.asia",
            "malware_url_count" : "19757",
            "last_seen" : "2013-10-24T02:14:12-07:00",
            "first_seen" : "2013-10-24T02:14:12-07:00"
         },
         {
            "url_count" : "1",
            "domain" : "moousi.com",
            "malware_url_count" : "0",
            "last_seen" : "2014-01-20T02:17:57-08:00",
            "first_seen" : "2014-01-20T02:17:57-08:00"
         }
      ],
      "risk_factor" : 100,
      "host" : "NXDOMAIN",
      "factoring" : [
         {
            "ipviking_category_factor" : 101,
            "geomatch_factor" : 0,
            "asn_threat_factor" : 2,
            "iana_allocation_factor" : "-2",
            "asn_record_factor" : "-2",
            "ipviking_geofilter_factor" : 0,
            "ipviking_geofilter_rule" : 0,
            "ipviking_personal_factor" : "-1",
            "geomatch_distance" : 0,
            "data_age_factor" : "30",
            "bgp_delegation_factor" : "-2",
            "country_risk_factor" : "4.1",
            "ip_resolve_factor" : "12",
            "search_volume_factor" : "0",
            "region_risk_factor" : "5"
         }
      ],
      "risk_desc" : "Extreme risk involved",
      "customID" : 0,
      "factor_entries" : 13,
      "transID" : 0
   }
}
```

2.2. IPQ Score

The IPQ score ranges are from **0** (no or low risk) to **100** (extreme risk) with precision of one decimal place. This value is used to provide a reference point to the Norse customer as to the acceptability of the transaction or communication in question. The IPQ is the value assigned by IPViking to reflect the actual analyzed aggregated behavior of the IP address. The foundation of this value is in the more than 1,500 factors used to evaluate the IP address at the time of submission.

2.3. Factoring

For each IPQ score returned, IPViking™ delivers a set of factors that represent a roll-up of the 1500 criteria used to analyze and create the IPQ score. Adding these factors together (they can run negative) will total to the IPQ score. Note: IPQ score will be capped in the range of 0-100.

The factors provide additional context to the categories, protocols and geolocation information returned with every IPQ. Policies can be created using these data points to control access to critical assets.

The next level of analysis is to understand which factors if any is the dominant weighting in the overall score. Generally, any factor score above 10 is more significant. Finding the largest absolute value is a quick way to identify the dominant factor(s).

In the example below, the IPViking Category Factor and Data Age Factor are the largest absolute factors, although note that there is some risk from the Country Risk Factor and IP Resolve Factor.

Factoring

Country Risk Factor	4.1	**Region Risk Factor**	5	**Ip Resolve Factor**	8
Asn Record Factor	-2	**Asn Threat Factor**	2	**Bgp Delegation Factor**	-2
Iana Allocation Factor	-2	**Ipviking Personal Factor**	-1	**Ipviking Category Factor**	101
Ipviking Geofilter Factor	0	**Ipviking Geofilter Rule**	0	**Data Age Factor**	30
Geomatch Distance	0	**Geomatch Factor**	0	**Search Volume Factor**	0

IPVIKING CATEGORY FACTOR

Contextual statistical analysis using statistical inference towards actual collected Internet intelligence

What causes it to rise?
Based on the Context Rationale field. Each category ID has a multitude of factors that comprise this value including time, weight of Category ID, frequency, etc.

Implication
This is a significant indicator to the risk score. Remediation rules should focus on the value of this factor.

DATA AGE FACTOR

Factor assessment derived from the factors timestamps used in the entire risk assessment

What causes it to rise?
Timeline algorithms, inference engine, category ID, longevity component of category ID, on/off aspect implies bots or mobile devices

Implication
The more recent an event, the higher the potential for the IPQ score. This is a complex factor that takes multiple sub-factors into consideration. This is a significant factor to be used in creating mitigation policies.

IP RESOLVE FACTOR

Factor assessment of the current and historical DNS reverse lookup for the submitted IP.

What causes it to rise?
IP won't resolve correctly or consistently; Resolves to a blacklisted IP; Timeline indicates it's resolving to different domains too often. While an inability to resolve can be attributed to a remote DNS server providing incorrect data, or other configuration issues, IPViking attempts to take these networking issues into consideration when calculating the factor.

Implication
IP may have been hijacked or carries additional risk due to not having reliable DNS entries.

ASN RECORD FACTOR

An assessment based on the presence of the AS (Autonomous System) the IP is assigned to is calculated.

What causes it to rise?
AS number to BGP relationship changes; this factor is calculated every 5 seconds.

Implication
This factor might be high when the ASN does not match a valid autonomous system, suggesting a forged request.

ASN THREAT FACTOR

Statistical analysis on current and historical incidents resolved to AS numbers assessment

What causes it to rise?

If number of attacks associated with hosts within this AS rises.

Implication
This factor might be high when the AS that this IP belongs to is commonly associated with high-risk activity. An IP might have a high score only because of its relationship to the company that owns the routing responsibility. This is, "Guilt through association". It is recommended that the user find a new ISP.

BGP Delegation Factor

BGP delegation records current and historical assessment of multiple characteristics.

What causes it to rise?
If IP is not in the BGP table or it changes value in a short period of time.

Implication
There is IP squatting, hijacked IP, or IP spoofing taking place.

IANA Allocation Factor

Assessment based on current and historical IANA allocation records for the IP submitted.

What causes it to rise?
IP unallocated or unassigned, private or router addresses used.

Implication
No allocation indicates this is not a legitimately assigned address. Indicates malicious traffic.

Country Risk Factor

This is the hourly calculation reflecting the country risk factor assessment.

What causes it to rise?
The number of hosts participating in an attack, the percentage of hosts in attacks currently emanating from this country, etc.

Implication
These countries are current or recent participants in a cyber attack.

Region Risk Factor

This is the statistical analysis reflecting the region risk factor assessment.

What causes it to rise?
This is the same as the country risk factor but at regional level.

Implication
These regions are current or recent participants in a cyber attack.

GEOMATCH FACTOR

Factor based on the submitted geographical address compared to the calculated IP address geographical location.

What causes it to rise?
An example could be that multiple transactions are attempted from different locations in a short period of time.

Implication
If the distance value is high but location is a hotel or near an airport, then score could be low. The factor takes all of these variables into consideration.

GEOMATCH DISTANCE

Assessment of the actual distance in miles used to derive the geomatch factor, derived from haversine or spherical law of cosines. This is useful when risk or fraud policies correlate highly to geolocation or changes in geolocation. As an example, in fraud when comparing a billing address with the actual geolocation of the IP submitting a transaction. This feature is optional and invoked in the API by passing additional parameters to the API call—details are provided in the Developer's Guide. A whitepaper is also available from Norse that further describes the geomatch algorithms.

What causes it to rise?
"As the crow flies" distance calculated between a billing address and the computed geographical location of a submitted IP.

Implication
Response is in miles. The greater the number, the greater the difference between the supplied billing address and the calculated geolocation. A large number may indicate fraud, but a small number can also indicate "friendly fraud" where a buyer is having remorse and trying to deny the validity of a purchase.

IPVIKING GEOFILTER FACTOR

Factor derived from the user manipulated GeoFiltering system. This optional system allows an IPViking API user to set ACL rules on good or bad geo-regions (country down to zipcode). These geo ACLs can be configured via the API or the portal. When an IP address's computed geo-location (by Norse) matches an ACL rule, its score is adjusted (+100 for a "bad" region and -50 for a "good" region).

What causes it to rise?
Matching of rules created in customer portal or via the API. See Developer Guide for more details.

Implication
User can create complex rules for isolating high-risk traffic when risk or fraud correlates to specific geographies of the source IP.

IPVIKING GEOFILTER RULE

Rule ID from user submitted rules used to derived the above factor

Implication
The rule ID that was triggered in the table is displayed.

SEARCH VOLUME FACTOR

The search volume factor applies some algorithms that are similar to "velocity" concepts within fraud and are most applicable to those use cases. It reflects a higher score when a higher risk IP address has been queried across customers at an elevated frequency/rate in the IPViking API, indicating that it has been both malicious and active in a broader context.

IPVIKING PERSONAL FACTOR

The IPViking Personal Factor is reserved for internal use, though may have a slight value in practice. It can be ignored for practical purposes.

2.4. Context / Rationale (Categories-Protocols)

IPViking is looking for "dark intelligence". This is where the bad actors interact using the same tools they use to put forth what is happening and more importantly, set up for future bad actor behavior. With Norse, we allow our customers to see into the future, with what bad actor behavior is coming towards them.

Protocols are sub-categories and are attached to a parent category, providing more detail in some cases. Some categories and protocols are provided for context for analysis or investigation, not because they significantly affect risk. In some cases, the category and protocol are general, indicating that Norse had general indicators of the type of activity but did not match to a more specific signature. This can occur with unidentified or new binaries.

A complete list of categories and protocols are listed in the Appendix.

The following are some of the most prominent categories that can affect the security posture of your business.

ANONYMOUS PROXIES

Norse only identifies anonymous proxy networks, not benign proxies such as corporate web proxies.

Anonymous proxies are tools that attempt to make activity on the Internet untraceable. It is generally a server computer that acts as an intermediary and privacy shield between a client computer and the rest of the Internet. It accesses the Internet on the user's behalf, protecting personal information by hiding the client computer's identifying information.

There is a significant risk associated with anyone trying to use an anonymizing proxy to conduct payment transactions. In the case of protocols associated with this overall proxy category IPViking is uniquely capable of identifying various types of proxies including both IP based proxies and TOR exit nodes (both known and hidden). TOR is used by criminal actors to hide their actual IP addresses and has both public and hidden exit nodes. IPViking provides direct view of over 99% of **all** TOR exit nodes on a global level, with its internal database holding ~200K Tor Exit nodes compared to 3-4K typically freely shared on the Internet. Additionally, investment has been made to discover new Tor exit nodes within several minutes.

In addition to Tor, Norse applies similar discovery to SOCKS4/5 proxy networks, as well as common Web VPN networks.

The category is Proxy, and the protocol can be:

- IP Proxy (SOCKS4/5)
- Tor Exit (Tor)
- Web Proxy (Web VPN)

BOGONS

Bogon is an informal name for an IP packet on the public Internet that claims to be from an area of the IP address space reserved, but not yet allocated or delegated by the Internet Assigned Numbers Authority (IANA) or a delegated Regional Internet Registry (RIR). The risk associated with a Bogon is that it very well could be a hijacked IP address that is being used to conduct the transaction.

The IPViking technology is capable of resolving these issues live, as the IP address is presented for analysis. Norse understands that not all Bogons are necessarily hijacked IP addresses. To ascertain the risk value associated with a IP address Norse conducts a normalization of the IP address using among other factors the ISP BGP tables, ARIN, IANA and RIR tables to insure that the IP address is operating within the correct and assigned network. The next level is to see if the IP address has been announced. The announcement

is the broadcast of the IP address to the router level global, announcing the IP address is now assigned to an organization. The next level is to see if that the IP address actually resolves to the correct assigned network.

In the case where an IP address is operating in the correct network but has not been announced by the ISP Norse attaches a Bogon Unadvertised category and protocol to the IP address. This among other variables used may increase the IPQ score. As each ISP has a different "roll rate" in which new IP addresses are announced this is taken into account in the analysis conducted. It is quite possible that a IP address classified as a Bogon Unadvertised with an initially high IPQ score can change within the time the ISP announces it. The IPQ scoring would automatically adjust based on the announcement.

In the case of an IP address observed operating in an incorrect network it would be assigned a category of Bogon Unassigned. This type of Bogon has a greater risk associated with it as it is not operating within its assigned ISP network. When the IPViking technology attempts to resolve the Bogon Unassigned IP address during the query and it fails to provide a response it is considered to be a hijacked IP address and that factor is part of the overall factoring that results in the IPQ score. In the case of a Bogon Unassigned IP address the score value would place it in a high-risk environment.

Bogon Unadvertised activity by itself does not significantly weight IPQ risk score, and is included for context for investigation or analytics.

The category and protocol in these two cases would be:

- Bogon Unadv.IP Unadvertised
- Bogon Unass.IP Unassigned

Bots

Norse's current definition of a bot is more general than strict malware, and will include behavior that is automated in any suspicious way and could include repeated active scanning that violates RFC protocols.

Botnets comprise computers whose security defenses have been breached and control ceded to a 3rd party. Each such compromised device, known as a "bot", is created when a computer is penetrated by software from a malware (malicious software) distribution. The controller of a botnet is able to direct the activities of these compromised computers through communication channels formed by standards-based network protocols such as IRC (Internet Relay Chat) and HTTP (Hypertext Transfer Protocol).

Currently the parties responsible for creating and managing bots have moved to a significantly more difficult method of communications protocols to protect their botnets. The botnet communication activities are now being seen using *Peer-to-Peer* protocols (P2P), as the ability for most systems to identify and detect P2P transactions is very limited giving the botnet significant advantage in detection protection. However, IPViking has a long standing ability to detect and identity IP addresses using P2P protocols including the type of communications and other features necessary to uncover botnet communications.

IPViking identifies both individual bots as well as those identified to being command and control bots used to manage individual bots.

The bot categories and protocols include:

- Botnet.Botnet CC (command and control)
- Botnet.Bot

Malware URL

This indicates that Norse crawlers have identified exploits being served or dropped at a specific URL on a web service. Both static and dynamic analysis has been applied to the payload.

Actual listing of the malware URLs and signature match from common AV databases are included in other context returned by the API:

Malware URLs

Category Name	MIME type	URL	Last Seen
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v1688/?vug&a...	2013-10-24T02:36:30-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:07-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00
Win.Adware.Agent-2573	binary	http://applicationscreditforally.asia/v686/?bl5&am...	2013-10-24T02:41:06-07:00

The category and protocol are:

- Malware.Malware URL

Passive DNS

This indicates that passive DNS records exist for this IP address. This does not significantly weight the IPQ score but is included as context for investigation or analytics.

The actual DNS records are further detailed in other areas of context returned by the API. A count of the malicious and total URLs discovered per domain are also returned:

Passive DNS

Domain	Malware / Total URLs	Last Seen
moousi.com	0 / 1	2014-01-20T02:17:57-08:00
applicationscreditforally.asia	19757 / 21840	2013-10-24T02:14:12-07:00
www.cuzzsoft.com	0 / 1	2013-10-02T09:42:37-07:00
cuzzsoft.com	0 / 1	2013-10-02T09:42:41-07:00
greatfilesarey.asia	140 / 174	2013-10-01T23:16:03-07:00

The category and protocol are:

- Passive DNS.Malware domain

HTTP

Various context of HTTP services in shared hosting sites is included. It does not significantly affect IPQ risk score. Often the type of content is identified in the protocol e.g. Explicit Content. This categorization is included for context.

TIMESTAMPS

There are two different timestamps presented in IPViking:

1. IPQ_Timestamp – This represents the time that the IP was submitted for scoring. The score is based on current conditions.
2. IPQ_Entries_n_last active – This is the time when this particular category was scored to this IP in our system. Norse maintains a 3-year history of category associations on every IP in the system. The "n" represents which category this timestamp references. An IP can have several categories associated with it over time.

Having the timestamps of the last time an IP with a malicious category was seen can be beneficial for a user interested in understanding trends and actual live threats. It is far more significant that a bot is active on an IP trying to connect to your infrastructure than if a bot was seen 3 weeks ago with no activity since then. The "date last seen" timestamp is a significant component when calculating risk.

2.5. Network Telemetry

Every IP is associated with some public facing infrastructure. Understanding the "health" of that system can be critical in evaluating whether or not an IP should be allowed to connect to your critical assets. IPViking returns information associated with the public infrastructure with scoring and information that can be used to make critical decisions within the business.

IP Address and Hosts

Addresses and hosts have a relationship with each other that is well understood. Like any good protocol, DNS attempts to use that relationship to help with people interacting with machines in a way that is both efficient and scalable. IP addresses resolve to hosts and visa versa. When this process doesn't work, we have the first level of concern that things are not as they seem. As the Internet is built on layers and relationships exist upstream and downstream from IP in question, IPViking returns information on this as it is all related to the risk of making a connection.

Service Providers/Autonomous System/Organization

The Internet is simply a network of networks. The hierarchy requires all IP traffic to pass through service providers (ISPs). When organizations sign up for Internet services they become part of the network and are susceptible to those security practices. Through no fault of their own, businesses and individual consumers are exposed to an insufficient understanding or concern for security practices and therefore need to have the information necessary to protect their assets and provide compliance where necessary.

An autonomous system is a collection of connected IP routing prefixes under the control of one or more network operators that presents a common, clearly defined routing policy to the Internet. Even though there may be multiple Autonomous Systems supported by an ISP, the Internet only sees the routing policy of the ISP. That ISP must have an officially registered Autonomous System Number (ASN). A unique ASN is allocated to each AS for use in BGP routing.

All ASNs are subject to assignment by IANA. Norse is a member of the IANA body that assigns and maintains the tables; access to the most current information available can be referenced as part of understanding current routing patterns.

2.6. Geographic Intelligence

Geolocation has had a longstanding benefit for systems that calculate risk. Norse has developed an organic method for calculating latitude and longitude coordinates with a high degree of accuracy. Current IPViking customers enhance their IPQ scores with this information.

2.7. Device Type

Norse identifies device information with attributes in the following areas:

- Device type: indicates general type of device. Possible values: desktop, ipad, mobile, server, appliance, laptop, unknown.
- Operating System: general operating system family (does not include flavors or versions). Example values: Linux, FreeBSD, Windows, Mac OSX
- Category: indicates general organizational or locational usage. Possible values: consumer, government, corporate, education, satellite, unknown
- Probability: reflects Norse confidence level of the device type information, the value range from 0 to 100 where 100 means 100% confidence.

3. APPLYING IPVIKING DATA

In some binary instances, using just the score is sufficient for determining risk. This can be the case in the payments industry where tests are done prior to submitting a credit card for bank authorization. If a predetermined threshold is exceeded, the transaction may be denied. For fraud detection, this has been a very accurate and successful measurement.

However, in the risk and fraud space, strictly using the IPQ score may not always provide enough context to make an allow/deny decision. For this reason, IPViking™ provides categories and other context, which can be used to further the decision making process.

Guidelines for applying IPViking data are discussed in this section.

3.1. Use Cases

FRAUD

Payments

In Norse's client base for eCommerce Card Not Present (CNP) transactions, a default score of **48** is the score used to define a non-acceptable risk value for the presented transaction and therefore is used to decline the transaction. This default score was derived over time using transactional scoring involving CNP merchants and millions of transactions measured against charge back data and returns which led Norse to recommend this as the most effective score value to base a decline on.

This threshold value is variable, and a customer may, based upon their review, change the declination (or block) level up or down based upon their own internal view of the data.

IPQ Score	Risk Color Mapping	Risk Name
0 ≤ ipq_risk_factor < 34	Green	Low
34 ≤ ipq_risk_factor < 54	Yellow	Medium
54 ≤ ipq_risk_factor < 89	Orange	High
89 ≤ ipq_risk_factor ≤ 100	Red	Extreme

A potential rule:

- ✓ An IPQ > 54
- ✓ IPQ > 40 & Malware activity within last month
- ✓ IPQ > 40 & Botnet activity within last month
- ✓ IPQ > 50 & Proxy evidence within last 3 months

Other

Other fraud scenarios exist outside of payments. In these cases, thresholding combined with categories and geofilter/geomatch rules are reasonable parts of the dataset to apply in policy decisions.

The actual thresholds will depend upon the exact customer, their business, and the traffic profile for fraud. Part of operationalizing policy would be to collect and analyze a statistically significant sample of data (e.g. one month) and based on that optimize the policies to result in the highest catch rates while minimizing false positives.

APPLICATION

Login

In an application context such as login, the IPQ score can be used as a top-level filter though context such as the category and protocol will be useful in providing more granularity.

For example, if malware compromise is a problem on the end-user's client machine, then a score > 60 along with recent (last 3 months) malware or bot activity, could result in requiring secondary authentication as well as a customer service interaction to help the customer perform an AV scan.

If brute-force attacks or reconnaissance is suspected, then a score > 60 along with Bot or Proxy activity could be used to introduce more detailed logging or an alert in the background for later analysis.

Other

Many other application scenarios exist, including mid-session analytics on user behavior. The actual factors or contexts will depend upon which are the most appropriate to the application context at the point in time.

If scraping of data is suspected, then any bot categories or higher-risk geo factors could be used to limit the user's session or force an email notification to the email address on file to verify the account.

It is not suggested that the actions be severe or binary (e.g. block, terminate session)—the advantage of integrating IPViking intelligence into the application, is that it provides an opportunity to take more nuanced and intelligent decisions when risk exists for the other party (via their IP).

In-line Network Security (Prevention)

For in-line points of integration where the IPViking data will be applied to taking action. Generally, the IPQ score can be used in a general sense to drive alerting or analytics.

For example, a range of 60-79 can be used for alerting or analytics on suspicious IPs that are worth looking at.

A stricter threshold of 80-100 can be used to drive more active measures such as severe alerts, escalated investigations or if desired, blocking.

Certainly, depending upon protocol and place in the network, some more evolved rules can be used. For HTTP traffic through a web firewall, the malware URL categories with a lower threshold (>50) may be worth alerting or correlating upon.

As with all in-line scenarios, it is advisable to take a measured approach to blocking and first detect and report.

SOC / NOC (Detection I)

As part of detection workflows and processes, the SOC or NOC may require simple guidance on how to use Norse intelligence. The data may be analyzed or correlated within a tool such as a SIEM or may be exposed more directly to the SOC / NOC engineer.

In both cases, the score is the simplest factor to use, and it is suggested that the Level I engineers focus on higher threshold with little chance for false positives e.g. IPQ score > 80.

Forensics and Incident Handling (Detection II)

For advanced security personnel doing investigations, the focus will be much more on the full context of the IPViking data, less on the score. In this case, the network telemetry (e.g. AS), the geolocation, and the categories and protocols (Bogon Unadvertised, Bot, Malware URL), and the passive DNS records all may provide useful context around an IP of interest.

Analytics (Cyber Teams)

For analytics, the full context should be stored in a datastore such as an RDBMS or home-grown platform (e.g. Hadoop). Allowing both relational and time-series queries into all attributes will provide value in looking at trends and changes as well as insight into high but lower risk IPs.

Correlating the Norse intelligence with other internal data is one approach to gain additional insight.

3.2. Protocol

A different view will be to look at Internet or application protocols involved with the IP address when correlating with Norse data. Whether in a prevention or detection scenario, knowing that HTTP traffic is the protocol in question, would lead a user to looking at categories specific to that protocol along with a lower threshold—as the combination would still flag relevant results.

Note that HTTP and general TCP/UDP traffic will be the most common protocols. Today, Norse data is not specifically targeted towards SMTP risk assessment or spam scenarios.

3.3. Point in the Network

The point in the network will also be another way to approach analysis of the data. Often this correlates to the scenario and protocol and whether it is a prevention or detection scenario. Here are common network locations:

- Perimeter
- IPS/IDS
- SIEM
- Host

Perimeter and in-line locations tend to be high-volume and have limited application context. In these cases a higher threshold of 80 is a good starting point for action or automation of any kind.

A SIEM aggregates a wide array of data from all over the network, here more evolved rules may apply that take into account user and task and may include categories and protocols for automation. Whereas investigative scenarios would look at all context to be interpreted and applied by an advanced security analyst.

Many of the Host scenarios would revolve around detection and forensics. Here it would be less common to take automated action, but instead provide full context to an investigator handling an incident.

4. OPERATIONS

4.1. Overview

This section is to be used as a guide and template in creating one more specifically suited to the particulars of your business.

4.2. Considerations

4.2.1 INTERNAL DEVICES

The Internal device communicating has a large impact on the context of the data. Knowing that the internal device is a web server vs. a smart TV gives a very different perspective. When rules are being followed by operations personnel it's important that they can easily identify what the different internal device types are so that thy can act accordingly. These are some examples of possible device type groups: Web server, database server, monitoring server, security appliance, networking infrastructure, point of sales system, employee laptop, printer, smart TV, domain server, name server, wifi / guest access.

How sensitive the information a device contains is also very important to understanding the context of the internal devices. More strict rules may be applied for protecting systems storing or with access to sensitive data.

This does not exclude the possibility that an internal device can be compromised and used in suspicious activity.

4.2.2 COMMUNICATION DIRECTION

Determine if it's outbound or inbound communications. This can be difficult with certain appliances depending on their network perspective. It can also be difficult given too narrow of a window of time. How important this is will largely depend on the internal device type and what is being communicated (FTP, SSH, HTTP,SMTP). For example: employee laptop traffic would normally be outbound to external resources. It would be unusual for an external resource to initiate a conversation to an employee laptop. A web server would have the opposite pattern. A variety of external resources will initiate contact with a web server. It may be unusual for that web server to initiate communication outbound, unless it was intentionally programmed to do so.

4.2.3 COMMUNICATION PROTOCOL

The information being communicated is very important to understanding the security implications. This level of detail may not be available at the time that something is brought to your attention. If the detail isn't there it could prompt further specific logging or monitoring to investigate further. The communication protocol typically refers to something that can identify the purpose of the traffic and its payload. SMTP, FTP, SCP, HTTP, DNS are a few examples. These should not be assumed based on a TCP port that is suggested for them. There is a lot of non-HTTP traffic that goes over port 80 simply because

port 80 is typically open on a firewall. A web server would be expect to have inbound HTTP communication, but inbound SCP or SSH may be very unusual.

4.2.4 Norse Intelligence on External Device

Norse IPViking provides a lot of security relevant intelligence to help you determine what actions to take based on all of the other considerations.

1. IPQ Score: A summary score, higher score = higher risk.
2. IPViking Categories: Detections of activity such as botnets, and HTTP categories.
3. Data Age: An indicator of how long it's been since a detection of an IPViking category was identified.
4. Device Type: description of a device type such as a server
5. Geo Factors: Risk based on geographic information such as other malicious activity in the region.
6. ASN / IANA Factors: If the IP is routable and advertised
7. Malware URLs: Identified URLs as malicious.
8. Passive DNS: Malicious domains registered on a name server

4.2.5 Expected Behavior

Similar to the points mentioned above, you set a profile per device type in relation to internal and external resources, and look for deviations. There are a lot of periodic events in normal operations. It can take a month or more to add the typical exceptions that would occur outside of a daily schedule.

You industry also comes into consideration for expected behavior. How widely distributed your business is, what you consider normal operations, and the sensitivity of the information you are guarding.

4.2.6 Exceptions

Some activity will fall outside of the larger rules. If a rule is working well for the group that it is designed around, creating exceptions can be the best way to prevent false positives without creating overly complex rules. When enough exceptions have enough common elements to create a new rule it's recommended to replace them to avoid confusion.

4.2.7 Action

After all rules are applied through systems or processes there are a limited set of actions that can ultimately be performed.

1. Allow: expected, benign or trusted
2. Block: suspicious, malicious, or unusual
3. Investigate and Monitor more closely: Undecided
4. Notify someone: Predefined contact
5. Modify a rule to alter future outcomes: Update procedures.
6. Adjust Score to re-evaluate in rules

4.3. The Living Document

Any set of procedures needs to be regularly reviewed and updated. Changing devices, expected behaviors, finding new exceptions can all have an impact on the rules in a system or operations manual.

4.4. How Norse Can Help

Norse can help guide you through this process. Any information that addresses the considerations mentioned in section 2 of this document will help if already assembled. With this information and enough traffic data to develop a profile, Norse Professional Services can get you started with a rule set tuned to your specific requirements.

4.5. Example Rules

These are example rules. They are not intended to be accurate for any specific business needs.

Internal	Dir	Protocol	Score	Age	Category	Action
Web	out	*	34+	*	*	Block
Employee	out	HTTP, SMTP	54<	*	If not malware or botnet	Allow
Employee	in	*	34+	*	*	Block
POS	in	*	48+	*	*	Block
*	*	*	34+	*	Countries outside of business deals	Block
Network	*	*	*	negative	Only bogon	Remove IP resolve factor from score

5. CAPABILITIES

Norse has a unique global data collection and analysis infrastructure for public IP traffic, with a focus on scoring risk and identifying risk factors from the dark side of the Internet, for each IPv4 address. We provides both products (SaaS) that use this risk information and API access to the risk information to mitigate network threats including eCommerce fraud, network attacks, compromised logins, and compromised internal hosts.

Norse's global distribution of sensors, placed to maximize coverage and sample rate of malicious or questionable Internet traffic, provide the current status of the external IP address space. Analysis of malicious or questionable traffic includes bad actors' use of: IRC, TOR, P2P, free open source services (DNS, SSH, VPN), unassigned or unadvertised address space via IANA and BGP information, attacks against Norse honeypots, Netflow analysis, and web crawling for specific terms and data of interest. The information from these multiple data sources is continuously analyzed and correlated across over 1500 criteria, resulting in a risk metric, contributing risk factors (rationale), and geolocation information for each IP address. Historical data of up to 2 years is also factored into the risk analysis.

The delivery of threat scores and corresponding factors is through a SaaS-based service via an API over HTTP/S, designed for integration into decision points in a network or application. Design of the Norse delivery infrastructure has focused on: comprehensive coverage of bad public IP traffic, high sample rate, and immediate delivery of current risk state and assessment.

APPENDIX A: CATEGORIES AND PROTOCOLS

Categories

Below are tables detailing the categories returned in the API request. There is a parent-child relationship between categories and protocols.

Category ID	Category Name
1	Explicit Content
2	Bogon Unadv
3	Bogon Unass
4	Proxy
5	Botnet
6	Financial
7	CyberTerrorism
8	Identity
9	BruteForce
10	Cyber Stalking
11	Arms
12	Drugs
13	Espionage
14	Music Piracy
15	Games piracy
16	Movie piracy
17	Publishing piracy
18	StockMarket

19	Hacked
20	Information piracy
21	High risk
22	HTTP
31	Malicious Site
41	Friendly Scanning
51	Web Attacks
61	DATA Harvesting
71	Global Whitelist
81	Malware
82	Passive DNS

Protocols

Below are tables detailing the protocols returned in the API request. There is a parent-child relationship between categories and protocols. Note that the protocol information includes a specific protocol name and an overall protocol.

Protocol ID	Protocol Name
7	eDonkey User
17	Ares User
19	Gnutella User
30	IP unassigned
31	IP unadvertised
32	Tor Exit
33	IP based proxy
34	Web Proxy
40	Botnet CC
41	Bot
50	Carding user
51	Carding Generator
52	Financial Fraudster
53	Financial publishing user
54	Insider Information Leak
55	Merchant Submitted fraud
60	Intelligence collector
70	Identity Phishing

80	Brute Force attacker
90	Cyber Stalker
100	Arms dealer site
110	Mule recruitment
111	Narcotics Soliciting
112	Pharmaceutical soliciting
120	Industrial Espionage Server
130	Hacked Computer
140	Pump and dump user
141	Short Scalping Computer
150	High risk user TMI
151	High risk Network
152	High risk site
153	Porn Content
156	Gambling
157	Chat
158	Web Radio
159	Webmail
160	Warez
161	Shopping
162	Advertisement
163	Movies
164	Violence

165	Music
166	Hacking
167	ISP
168	Drugs
169	Agressive
170	News
171	Redirector
172	Spyware
173	Dating
174	Dynamic
175	Job Search
176	Tracker
177	Models
178	Forum
179	Web TV
180	Downloads
181	Ring Tones
182	Search Engine
183	Social Net
184	Update Sites
185	Weapons
186	Web Phone
187	Religion

188	Image Hosting
189	Podcast
190	Hospitals
191	Military
192	Politics
193	Remote control
194	Fortune Telling
195	Library
196	Cost Traps
197	Homestyle
198	Government
199	Alcohol
200	Radio TV
201	Zeus Bot
211	Butterflies Bot
221	Keylogger Bot
231	Zeroaccess Bot
241	Palevo Bot
251	ICE XX Bot
261	SpyEye Bot
271	DriveByMalware
281	Binary Site
291	DDOS scatter

301	Port Scan Scatter
311	SPAM Scatter
321	Iframe Hidden
331	Iframe Injected
341	.htaccess redirect
351	Bad Javascript
361	Phishing Site
371	Friendly Port scan
381	Manual WL entry
391	VPN AnchorFree.com
392	Malware URL
393	Malware domain