

From: [REDACTED] </O=ITHAKA/OU=FIRST ADMINISTRATIVE GROUP/CN=RECIPIENTS/CN=[REDACTED]>
Sent: Wednesday, September 29, 2010 5:13 PM
To: '[REDACTED]' <[REDACTED]@MIT.EDU>
Cc: '[REDACTED]' <[REDACTED]@mit.edu>
Subject: RE: JSTOR Abuse at MIT: All IPs Blocked

Hi [REDACTED],

Yes, I was glad to be mistaken about the Class A range. That said, note that this was an extreme case. We typically suspend just one individual IP at a time and do that relatively infrequently (perhaps 6 on a busy day, from 7000+ institutional subscribers). In this case, we saw a performance hit on the live site, which I have only seen about 3 or 4 times in my 5 years here.

The pattern used was to create a new session for each PDF download or every few, which was terribly efficient, but not terribly subtle. In the end, we saw over 200K sessions in one hour's time during the peak. We have implemented some multiple session creation limits that will not be affected even by our heaviest users, but will catch this kind of extreme case and truncate the activity much sooner going forward.

That said, I can confirm that we will continue to take the least intrusive and most measured approach to addressing excessive use whenever possible, and a Class C range is the most we have blocked at once in my tenure. I have no reason to believe we would need to be more heavy handed than that, but, of course it is possible that it might be warranted at some point. I highly doubt it.

Thanks,

[REDACTED]

-----Original Message-----

From: [REDACTED] [mailto:[REDACTED]@MIT.EDU]
Sent: Wednesday, September 29, 2010 4:24 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: JSTOR Abuse at MIT: All IPs Blocked

[REDACTED], thanks very much. We appreciate it.

Looking to the future, would it be possible to clarify that JSTOR will follow the protocol that was ultimately used here, shutting down not the class A range but the class C range, should an excessive use case emerge again?

We have not had a history of excessive use of JSTOR content from MIT, so the problems do not seem at this time to be widespread. In addition, we are finding that the industry norm at this time is shut down of the specific offending IP.

We can see that in some cases (as here, where the initial suspension did not stop the misuse) moving to suspend the class C range is a reasonable response. But it would be very helpful on our end if we could work out an agreement that a shutdown of the class A range will not be part of the standard initial response protocol for excessive use cases.

We look forward to your thoughts --

[REDACTED]

[REDACTED]

[REDACTED]

MIT Libraries

P [REDACTED]

[REDACTED]@mit.edu

<http://libraries.mit.edu/scholarly>

-----Original Message-----

From: [REDACTED] [mailto:[REDACTED]@ithaka.org]

Sent: Wednesday, September 29, 2010 4:03 PM

To: [REDACTED]

Cc: [REDACTED]

Subject: RE: JSTOR Abuse at MIT: All IPs Blocked

Dear [REDACTED],

Thank you for your reply. I will ask our [REDACTED] staff to reinstate the suspended Class C range as soon as possible and will confirm once completed.

Thanks,

[REDACTED]

-----Original Message-----

From: [REDACTED] [mailto:[REDACTED]@MIT.EDU]

Sent: Wednesday, September 29, 2010 4:01 PM

To: [REDACTED]

Cc: [REDACTED]

Subject: FW: JSTOR Abuse at MIT: All IPs Blocked

Importance: High

Hello [REDACTED],

We've investigated this case and, because the origin of the activity was a guest visiting MIT, we believe it will not recur.

We hope you will be able to restore the class C range that has been suspended based on this information.

Thank you,

██████████
██████████
██████████
MIT Libraries

P ██████████
██████████@mit.edu

<http://libraries.mit.edu/scholarly>

From: ██████████@jstor.org]

Sent: Sunday, September 26, 2010 12:31 PM

To: ██████████

Subject: JSTOR Abuse at MIT: All IPs Blocked

Dear ██████████,

I am writing you this afternoon to let you know that we have been forced to block access to JSTOR from MIT. Yesterday, around 6pm, we began to see hundreds of PDF downloads per minute occurring from multiple sessions at 18.55.6.215. As these requests began to affect performance of the public site, we were forced to deny access to this IP. Requests continued to pour in from this IP for some time, but were denied access. This clearly indicates robotic harvesting of PDFs which violates our Terms & Conditions of Use.

This morning, at around 8am, this activity started again from IP 18.55.6.216, forcing us to restrict access to the entire range of MIT IP addresses. We rarely take this level of response to abusive activity, but felt it necessary to maintain the stability of the web site for other institutions and users.

Once you have identified the responsible party and can assure us that this activity will not continue, we will be happy to restore access as soon as possible. That said, please note that the block had to be executed at the firewall level to prevent performance degradation (even after denying the PDF downloads, the requests themselves were so frequent it continued to be problematic) and thus will require coordination of our systems administrators to restore, which may take some time.

Please do let me know if I can be of additional assistance or if I can provide additional information. Having worked in this area for some time, I am well aware that this activity is normally a compromised username and password or a student/researcher unaware of the impact of their activities or that this method of gathering PDFs is in violation of our Terms and Conditions of Use. We routinely work with researchers through our dfr.jstor.org site or by providing data exports to accomplish the intended research aim and would be happy to do so in this case as well if that turns out to be the motivation.

Best,

██████████
██████████
JSTOR

██████████@<mailto:██████████@ithaka.org>ithaka.org

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. JSTOR is part of ITHAKA, a not-for-profit organization that helps the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways.