NETWORK INFORMATION HIDING

CH. 7B: COUNTERMEASURES FOR DISTRIBUTED COVERT CHANNELS

Prof. Dr. Steffen Wendzel

https://www.wendzel.de

Some Notes

- There are a couple of sophisticated countermeasures [1]:

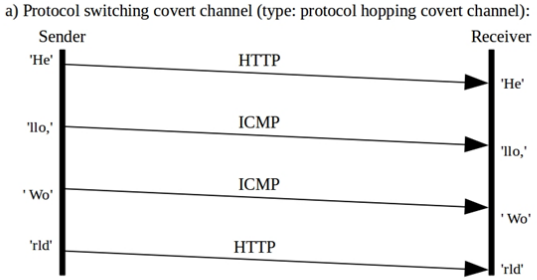
formation transfer. Active wardens with *active mapping* capability reduce the problem of ambiguities in network traffic (i.e. data that can be interpreted in multiple ways [LLC07]) by mapping a network and its policies [Sha02]. Afterwards, the mapped information is used by a NIDS to provide unambiguity [LLC07]. Based on the idea of active mapping and traffic normalization, Lewandowski et al. presented another technique called *network-aware active wardens* [LLC07]. Network-aware active wardens have knowledge about the network topology and implement a stateful traffic inspection with a focus on covert channel elimination [LLC06, LLC07].

[1] Wendzel, Steffen: The Problem of Traffic Normalization Within a Covert Channel's Network Environment Learning Phase, in Proc. *Sicherheit*, GI, 2012.

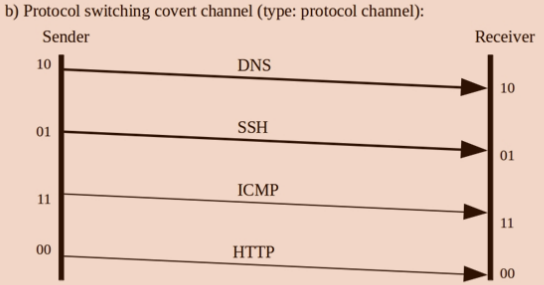
1. Protocol-channel Aware Active Warden

Remember Chapter 6?

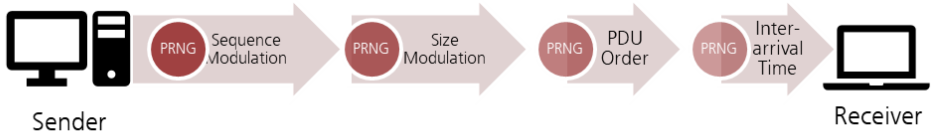
Protocol Hopping Covert Channel (PHCC) [2]:
Secret information is split over multiple network protocols to increase hurdles for a forensic traffic analysis.



Protocol (Switching Covert) Channel (PSCC) [1]:
Secret information is represented by the protocol itself.



Pattern Hopping [3]:
For every new piece of secret information a PRNG selects one of the patterns (+variation) to transfer the data.



[1] S. Wendzel, S. Zander: Detecting protocol switching covert channels, Proc. *Local Computer Networks (LCN), 2012 IEEE 37th Conference on*. IEEE, 2012.
[2] S. Wendzel, J. Keller: Low-attention forwarding for mobile network covert channels, Proc. Communications and Multimedia Security (CMS), 2011.
[3] S. Wendzel, S. Zander, B. Fechner, C. Herdin: Pattern-based Survey and Taxonomy for Network Covert Channels, ACM CSUR, Vol. 47(3), 2015.

Protocol Channel-aware Active Warden (PCAW)

- Protocol (Switching Covert) Channel-aware Active Wardens [1]
  - Limits bitrate of protocol switching covert channels
  - E.g. ICMP=„0“, UDP=„1“, Message „0101“ would then be represented by four packets with the order ICMP-UDP-ICMP-UDP

- How does it work?
  - Having a gateway that introduces delays for packets if they contain a protocol different from the previous one.

- State-holding: cache the last recently used protocol for each tuple (sender, receiver)
- Buffer: cache all packets that must be delayed

[1] S. Wendzel, J. Keller: Preventing Protocol Switching Covert Channels, in: Int. Journal Adv. Security, 2012.

PCAW: Example

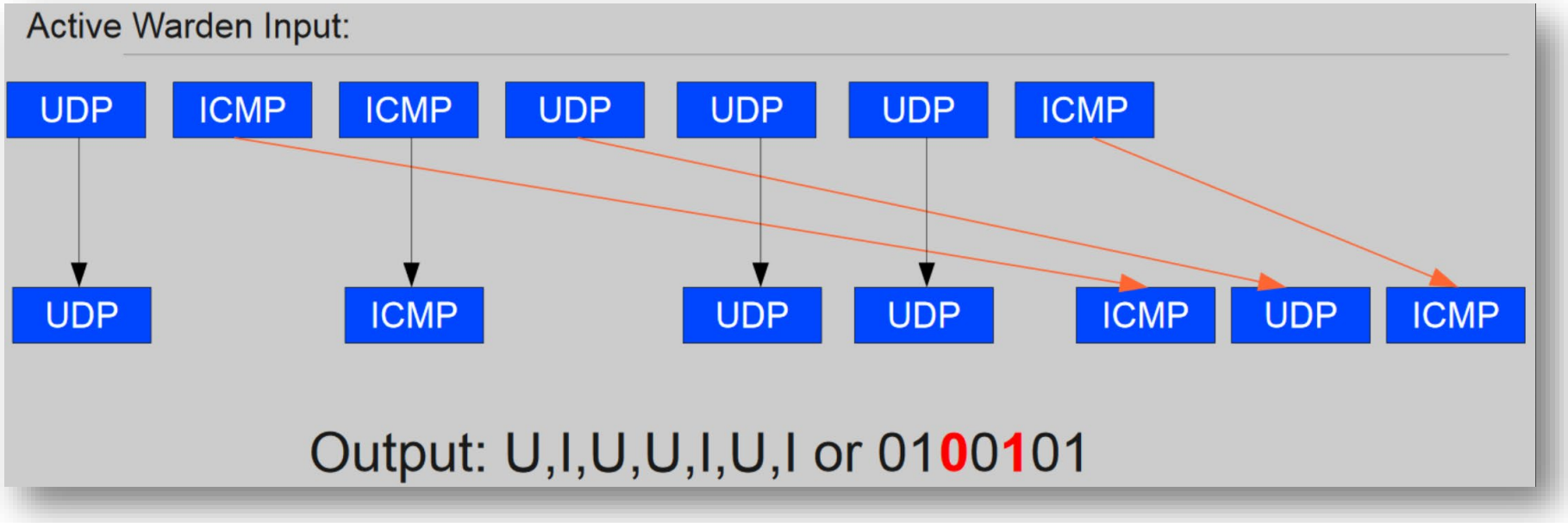


Fig.: S. Wendzel, J. Keller: Preventing Protocol Switching Covert Channels, in: Int. Journal Adv. Security, 2012.

PCAW: Impact on Bitrate [1]

- Bitrate calculation for a classical network CC (Tsai/Gligor):

$$B = b \cdot (T_R + T_S + 2T_{CS})^{-1}$$

b : number of bits to be transferred per transmission
T_x: time it takes to receive (R), send (S) and process (CS) the covert data.

- For a PSCC that utilizes n protocols, $b = \log_2 n$.

[1] S. Wendzel, J. Keller: Preventing Protocol Switching Covert Channels, in: Int. Journal Adv. Security, 2012.
[2] C.-R. Tsai and V. D. Gligor: A bandwidth computation model for covert storage channels and its applications, in Proc. IEEE Conf. on Security and Privacy, 1988, pp. 108–121.

PCAW: Impact on Bitrate [1]

- Given that we delay packets by *d* and that a switch of a protocol is taking place with probability *p*, while the transfer and covert data processing takes time *T*, we can modify the previous formula:

$$B = \log_2(n) \cdot (pd + T)^{-1}$$

- For a uniform coding with random input using *n* protocols, we know that $p = 1 - 1/n$. So, we can assume that:

$$B = \log_2(n) \cdot \left(\left(1 - \frac{1}{n}\right) d + T\right)^{-1}$$

- Other variants of PSCC exist where p and b may differ (see paper for details).

[1] S. Wendzel, J. Keller: Preventing Protocol Switching Covert Channels, in: Int. Journal Adv. Security, 2012.

PCAW – Expected Results

For a typical PSCC, we can reduce B to less than *1 bit/s* if we apply *d=2sec* for realistic T.

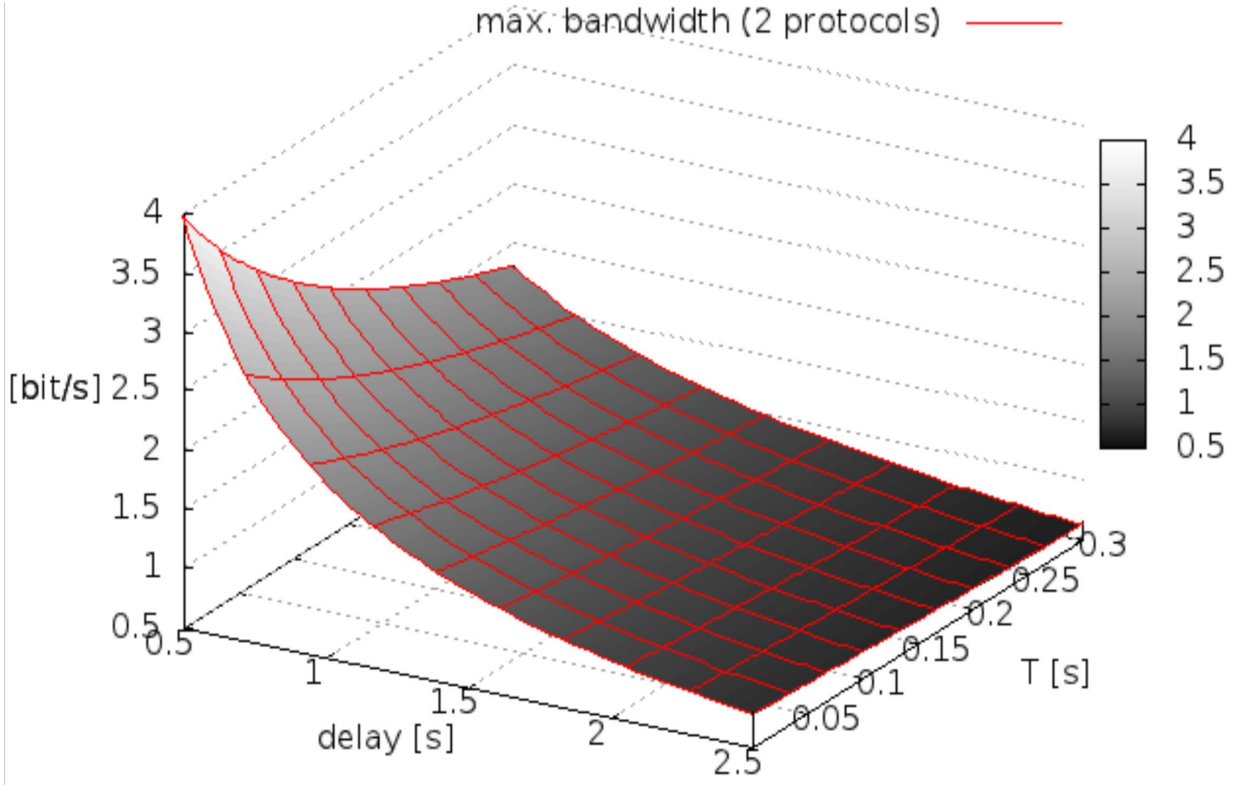


Fig.: S. Wendzel, J. Keller: Design and Implementation of an Active Warden Addressing Protocol Switching Covert Channels, in Proc. ICIMP, Iaria, 2012.

Randomized PCAW [1]

- Problem:
  - Attacker could try to determine d and then adjust own sending behavior, so that the warden becomes less effective.
- Solution:
  - Use a randomized delay, so that $d_r \in [0; d[$.
  - Turns out to provide excellent results.

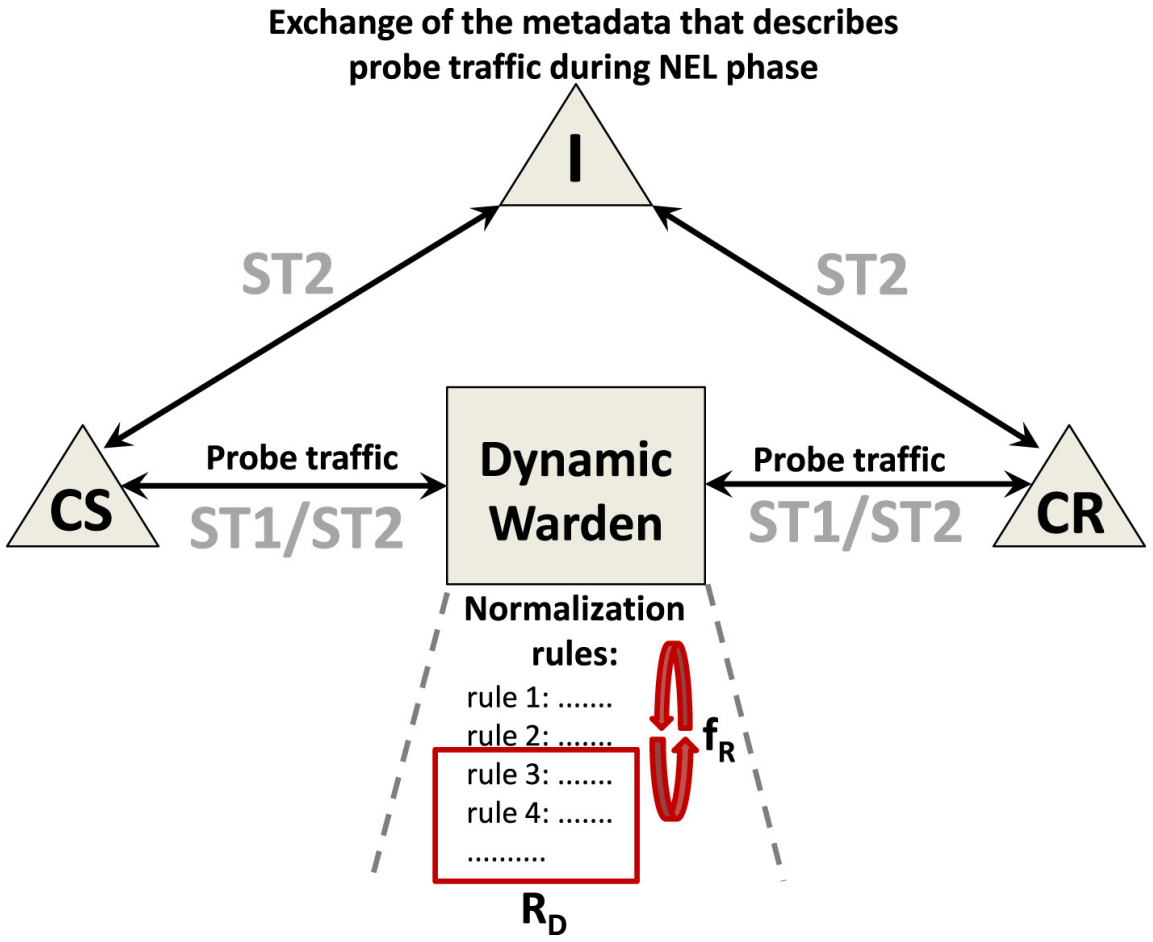
[1] S. Wendzel, J. Keller: Preventing Protocol Switching Covert Channels, in: Int. Journal Adv. Security, 2012.

2. Dynamic Warden

Dynamic Wardens

- **NEL+**-capable covert channel tools determine blocked covert channels; they can easily circumvent filers (see Ch. 6).

- Solution: Introducing a "Dynamic Warden".

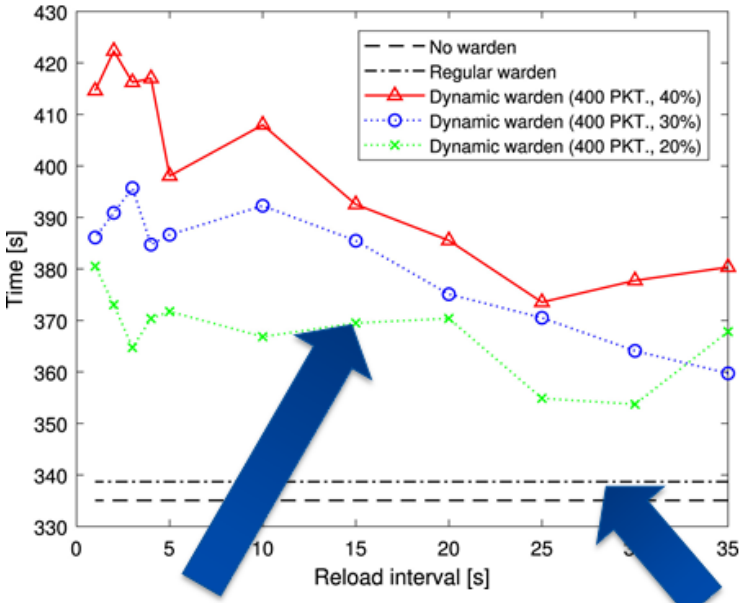


W. Mazurczyk, S. Wendzel et al.: Countering adaptive network covert communication with dynamic wardens, Future Generation Computer Systems, Vol. 94, pp. 712-725, Elsevier, 2019.

Dynamic Wardens: Results

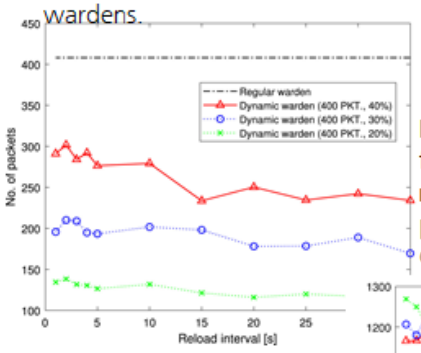
=> Results obtained from static configuration; each test repeated 20 times. Figures show average results.

Influence of the reload frequency on the **time needed to complete the transfer of 400 covert packets** for different types of wardens.

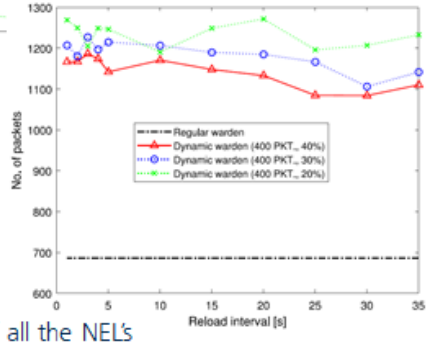


Dynamic warden with only 20% rule-set.

Influence of the reload frequency the on **the number of *normalized* packets** from CS to CR for different types of wardens.



Influence of the reload frequency the on **the number of *forwarded* packets** from CS to CR for different types of wardens.

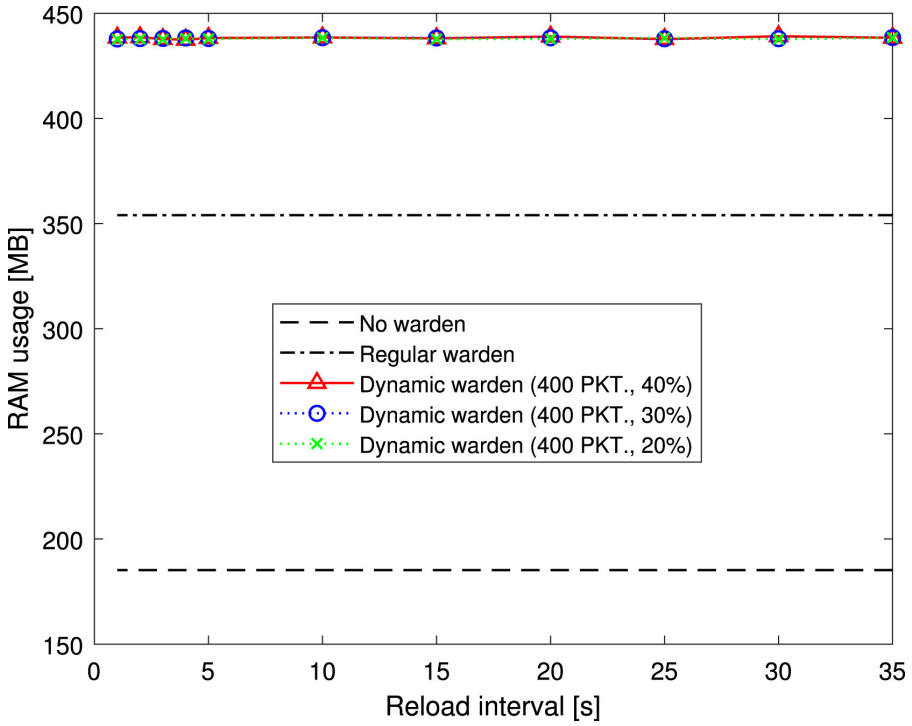


Regular warden with 80% rule-set (i.e. 80% of all the NEL's covert channels can be blocked)!

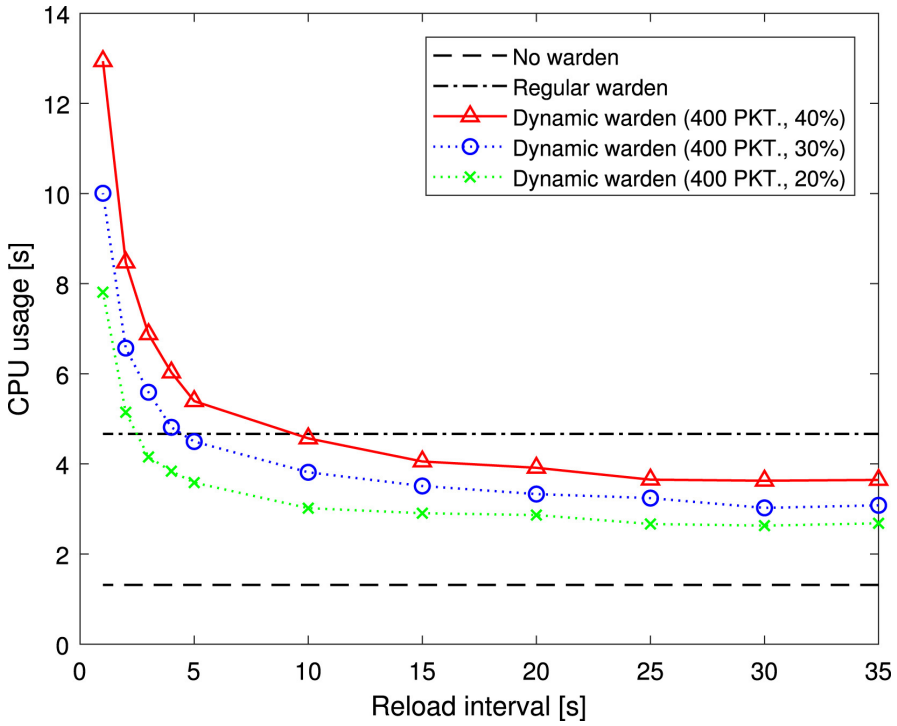
W. Mazurczyk, S. Wendzel et al.: Countering adaptive network covert communication with dynamic wardens, Future Generation Computer Systems, Vol. 94, pp. 712-725, Elsevier, 2019.

Dynamic Wardens: Results

Influence of the reload frequency on the **RAM usage** for different types of wardens (all wardens based on same Python code basis).



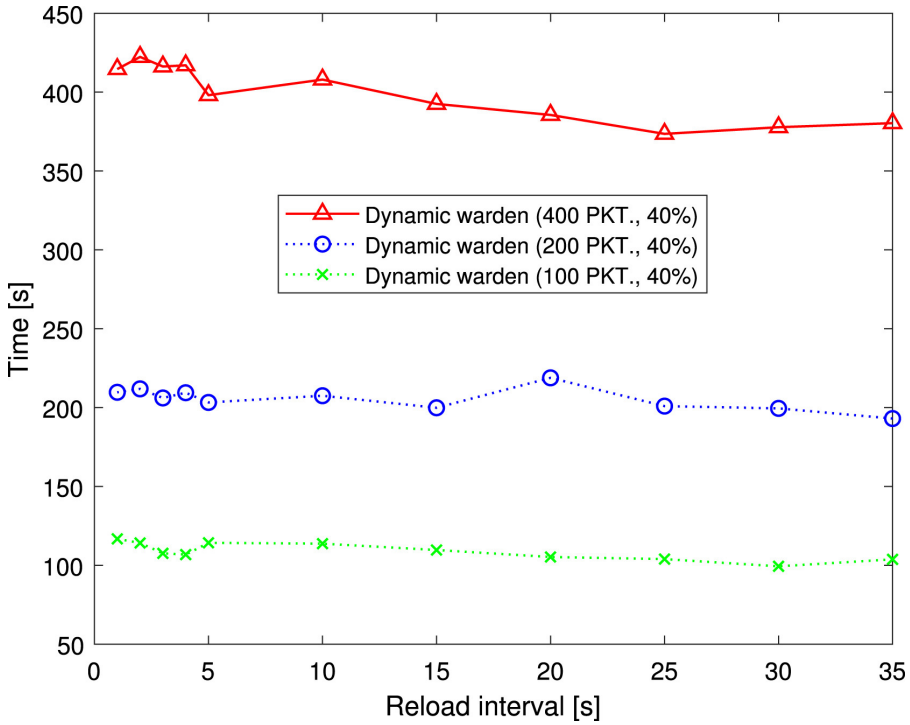
Influence of the reload frequency the on the **CPU usage** from CS to CR for different types of wardens.



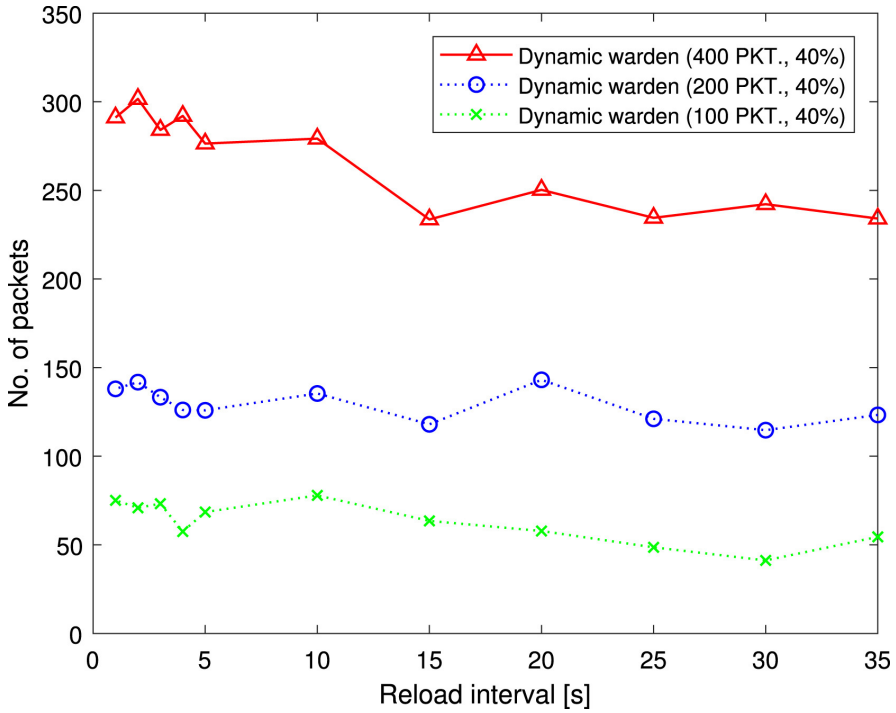
W. Mazurczyk, S. Wendzel et al.: Countering adaptive network covert communication with dynamic wardens, Future Generation Computer Systems, Vol. 94, pp. 712-725, Elsevier, 2019.

Dynamic Wardens: Results

Influence of the reload frequency on the **time needed to complete the transfer** of covert packets **for different lengths of the covert transmissions** (R_D=40%).



Influence of the reload frequency the on **the number of normalized packets** of covert packets **for different lengths of the covert transmissions** (R_D=40%).



W. Mazurczyk, S. Wendzel et al.: Countering adaptive network covert communication with dynamic wardens, Future Generation Computer Systems, Vol. 94, pp. 712-725, Elsevier, 2019.

Randomized Dynamic Warden: Results

Is it possible to load less filter rules on average by randomizing the number of loaded rules and the reload frequency?

- V1: $f_R \in \langle 1\text{ s}; 35\text{ s}\rangle$ and $R_D \in \langle 2\%; 100\%\rangle$ (i.e. between 1 and 50 rules). This means that the reload interval and the size of an active ruleset are selected randomly for the typical values investigated for the dynamic warden in the previous experiments.
- V2: $f_R \in \langle 1\text{ s}; 35\text{ s}\rangle$ and $R_D \in \langle 20\%; 40\%\rangle$ (i.e. the size of the active ruleset is randomly selected between 10 and 20 rules). Such values were tested for the dynamic warden in the previous sections.
- V3: $f_R \in \langle 1\text{ s}; 10\text{ s}\rangle$ and $R_D \in \langle 20\%; 100\%\rangle$ (i.e. between 10 and 50 rules). This means that the reload interval is selected from the values for which the best results have been achieved for the dynamic warden in the previous experiments.
- V4: $f_R \in \langle 1\text{ s}; 10\text{ s}\rangle$ and $R_D \in \langle 20\%; 40\%\rangle$ (i.e. between 10 and 20 rules) – both the reload interval and the size of the active ruleset are selected randomly in the ranges for which the best experimental results have been obtained for the dynamic warden investigated in the previous experiments.

Variant V3 offers the best results in terms of the

- **time needed to complete the covert transfer** and
- the **volume of traffic generated by the adaptive covert channel** parties (which is comparable with the best results obtained by the static setup for the dyn. warden)
- While offering **lower CPU and RAM consumption** (than static setup for the dyn. warden).

W. Mazurczyk, S. Wendzel et al.: Countering adaptive network covert communication with dynamic wardens, Future Generation Computer Systems, Vol. 94, pp. 712-725, Elsevier, 2019.

Dynamic Warden: Impact on Regular Traffic

- Impact of the dynamic warden on regular traffic is moderate in comparison to a regular warden.

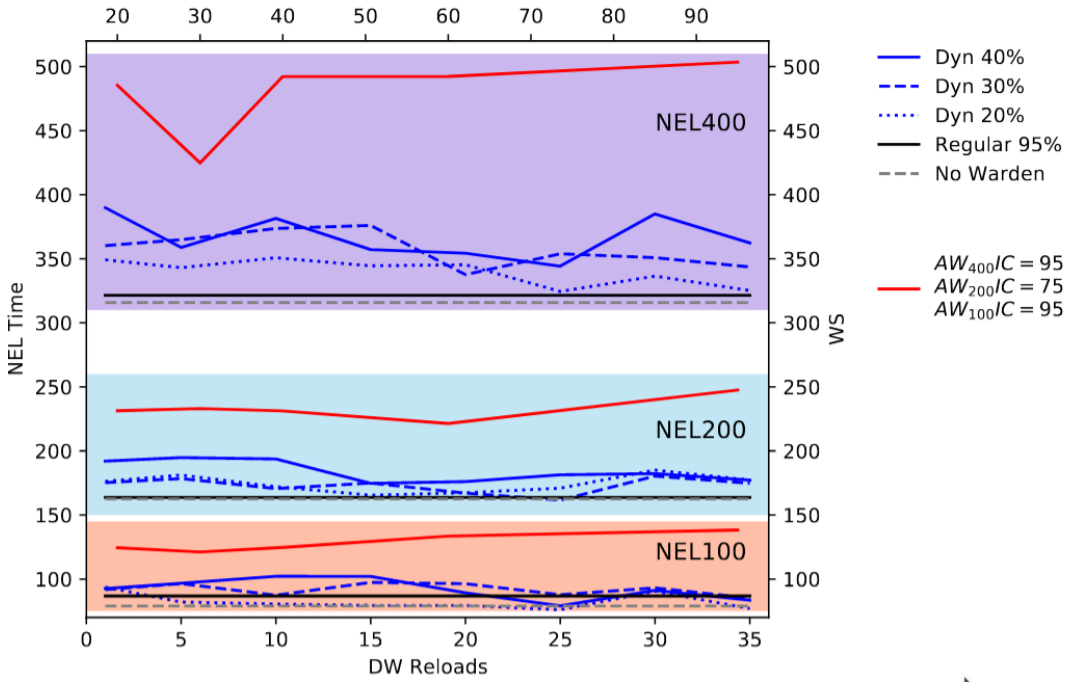
- Download of 100 MB file via HTTP with a static warden (95% rules) and a dynamic warden (40% rules).
  - Dynamic warden scenario: 7% shorter download time.

W. Mazurczyk, S. Wendzel et al.: Countering adaptive network covert communication with dynamic wardens, Future Generation Computer Systems, Vol. 94, pp. 712-725, Elsevier, 2019.

Dynamic Warden: Adaptive Version – slide not relevant for exam

Current work:

- Adaptive warden, i.e. a dynamic warden that adjust its behavior to the observed covert channel behavior.
  - implementation is ready
  - applying several rule selection algorithms for OS page replacement (LRU, …) to determine best strategy.



M. Chourib, S. Wendzel, W. Mazurczyk: Adaptive Warden Strategy for Countering Network Covert Storage Channels, in Proc. 46th IEEE Conference on Local Computer Networks (LCN), IEEE, 2021. https://doi.org/10.48550/arXiv.2111.03310