

A REVISED TAXONOMY OF STEGANOGRAPHY EMBEDDING PATTERNS

Steffen Wendzel^{1,2}, Luca Caviglione³, Wojciech Mazurczyk⁴, Aleksandra Mileva⁵, Jana Dittmann⁶,
Christian Krätzer⁶, Kevin Lamshöft⁶, Claus Vielhauer^{7,6}, Laura Hartmann^{1,2}, Jörg Keller², Tom Neubert^{7,6}

¹ Worms University of Applied Sciences, Worms Germany

³ National Research Council of Italy (CNR), Genova, Italy

⁵ University Goce Delcev, Stip, North Macedonia

⁷ Brandenburg University of Applied Sciences, Brandenburg, Germany

² FernUniversität in Hagen, Hagen, Germany

⁴ Warsaw University of Technology, Warsaw, Poland

⁶ University of Magdeburg, Magdeburg, Germany

Background and State of (Network) Steganography

HIDING PATTERNS

Hiding Patterns [1]

“Hiding Patterns describe the key idea of hiding techniques. They are kept on an abstract, non-detailed level, help cleaning up terminology, and can form a taxonomy.”

[1] S. Wendzel, S. Zander et al.: [Pattern-based Survey of Network Covert Channel Techniques](#), ACM CSUR, 47(3), 2015.

Comments on Patterns

- A technique can only be a pattern **if it occurs multiple times**. In general, the scientific patterns community agrees on a minimal number of three occurrences.
- **Pattern collections** comprise patterns of a given domain. They can be understood as **pattern catalogs*** (but the latter is additionally searchable, e.g. by an index of patterns).
 - e.g., a collection of user interface patterns
 - Problematic aspect: the link-ability of patterns between collections differs due to non-unified structures in which the patterns are described.

* Terminology not unified in the literature. We can agree on **collection==catalog** for this lecture.

- **Pattern languages** were introduced to solve the mentioned problems of pattern collections:
 - they provide a unified description for patterns
 - allow to build links/hierarchies between patterns
 - introduce aliases to prevent redundancies
- **PLML** (Pattern Language Markup Language, pronounced “Pell-Mell” [1]) is one dominating example of a pattern language.
 - We only use a subset of PLML as it suffices for our purposes.

[1] <https://www.cs.kent.ac.uk/people/staff/saf/patterns/plml.html>

- PLML allows the description of patterns (e.g. in XML).
- Hiding patterns can utilize various elements (attributes) of PLML/1.1:

Pattern Identifier	Name
Alias	Illustration
Description of the Problem	Description of the Context
Description of the Solution	Forces
Synopsis	Diagram
Evidence	Confidence
Literature	Implementation
Related Patterns	Pattern Links
Management Information	

* Newer version of PLML is available but the basic attributes remain. Not all attributes of the table above were used (+necessary) to describe hiding patterns.

The following PLML attributes were used

Table I. Used PLML/1.1 Attributes

Tag	Description
<pattern id>	Identifies a pattern within the particular catalog.
<name>	A correct assignment of a name for each pattern is important for the retrieval of a pattern when the pattern becomes part of a second catalog.
<alias>	Patterns can have different names, which are specified in the <alias> tag. The alias tag helps to find the same pattern when the pattern has different names in different catalogs.
<illustration>	An application scenario for the pattern.
<context>	Specifies the situations to which the pattern can be applied.
<solution>	Describes the solution for a problem to which the pattern can be applied. The attributes <i>problem</i> and <i>context</i> (cf. Fig. 1) are usually blurred but often not separated into two attributes.
<evidence>	Contains additional details about the pattern and its design. Moreover, the tag can contain examples for known uses of the pattern.
<literature>	Lists references to publications related to the pattern.
<implementation>	Introduces existing implementations, code fragments or implementational.

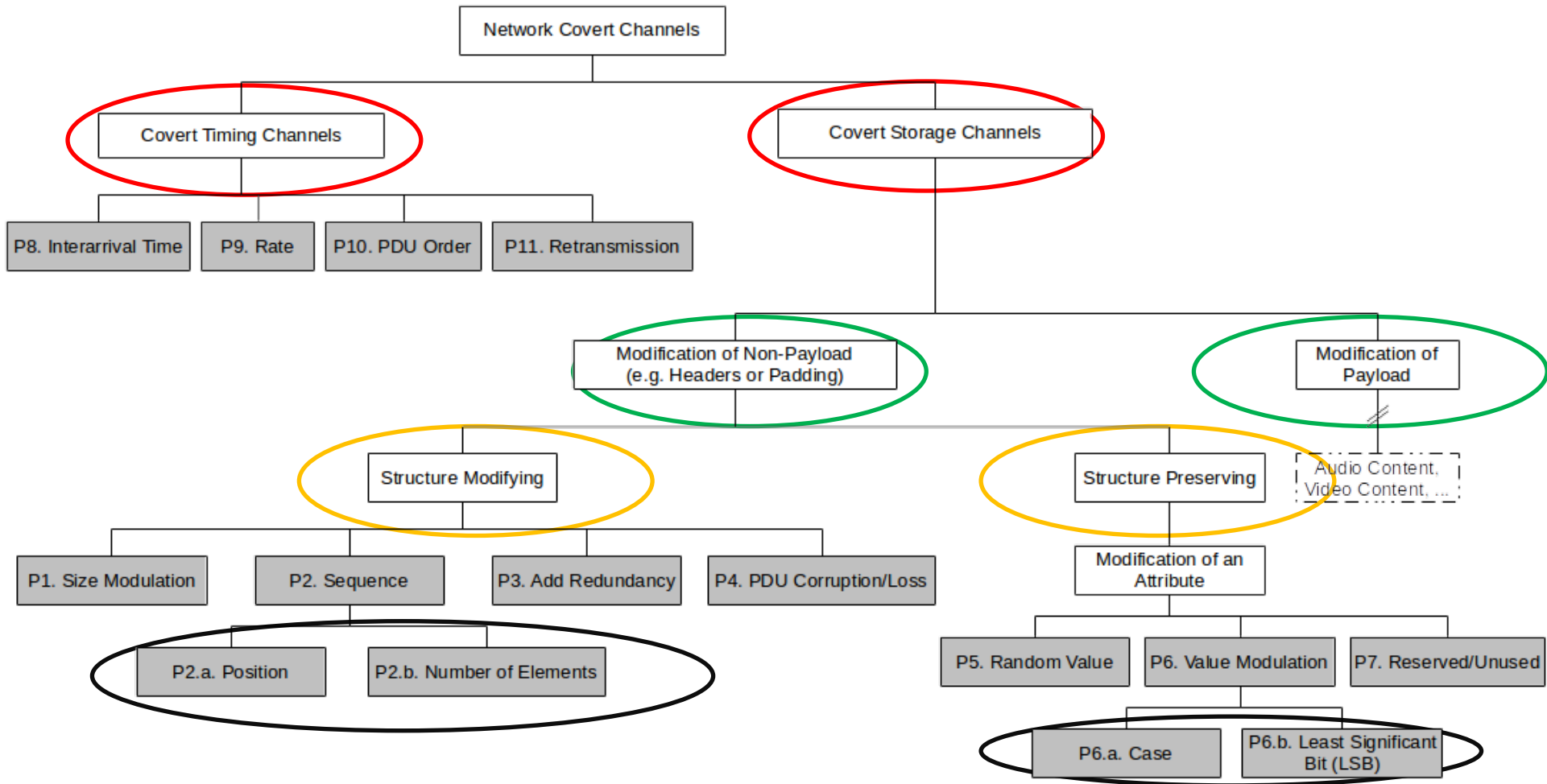
Fig.: S. Wendzel, S. Zander et al.: [Pattern-based Survey and Taxonomy for Network Covert Channels](#), ACM CSUR, Vol. 47(3), 2015.

Patterns in Network Information Hiding

- Approx. 170 network hiding techniques are known; they hide secret information in meta data of network traffic.
 - Inconsistent terminology.
 - Re-inventions very common.
- Instead of dealing with all these hiding techniques separately, we only need to understand the few hiding patterns.
- Initially **eleven** (later a more) patterns were found to describe all analyzed hiding techniques published since 1987.
- Also, patterns provide better taxonomies due to their several features (links and child patterns, alias handling, unified attributes, ...).

Patterns in Network Information Hiding [1]

Patterns were set in relation to other patterns to introduce a **new taxonomy** of patterns. The 109 hiding techniques could be described by only 11 patterns.



[1] S. Wendzel et al.: [Pattern-based Survey and Taxonomy for Network Covert Channels](#), ACM CSUR, Vol. 47(3), 2015.

Latest Version of Pattern Taxonomy

Currently, approx. 180 hiding techniques are categorized into 9 timing and 10 storage patterns, plus sub-patterns. Pattern names and their numbers were updated and extended in 2016, 2018, 2019, 2020 and 2021.

Based on:

S. Wendzel, S. Zander, B. Fechner, C. Herdin: [Pattern-based Survey and Taxonomy for Network Covert Channels](#), ACM CSUR, Vol. 47(3), 2015.

Extended by:

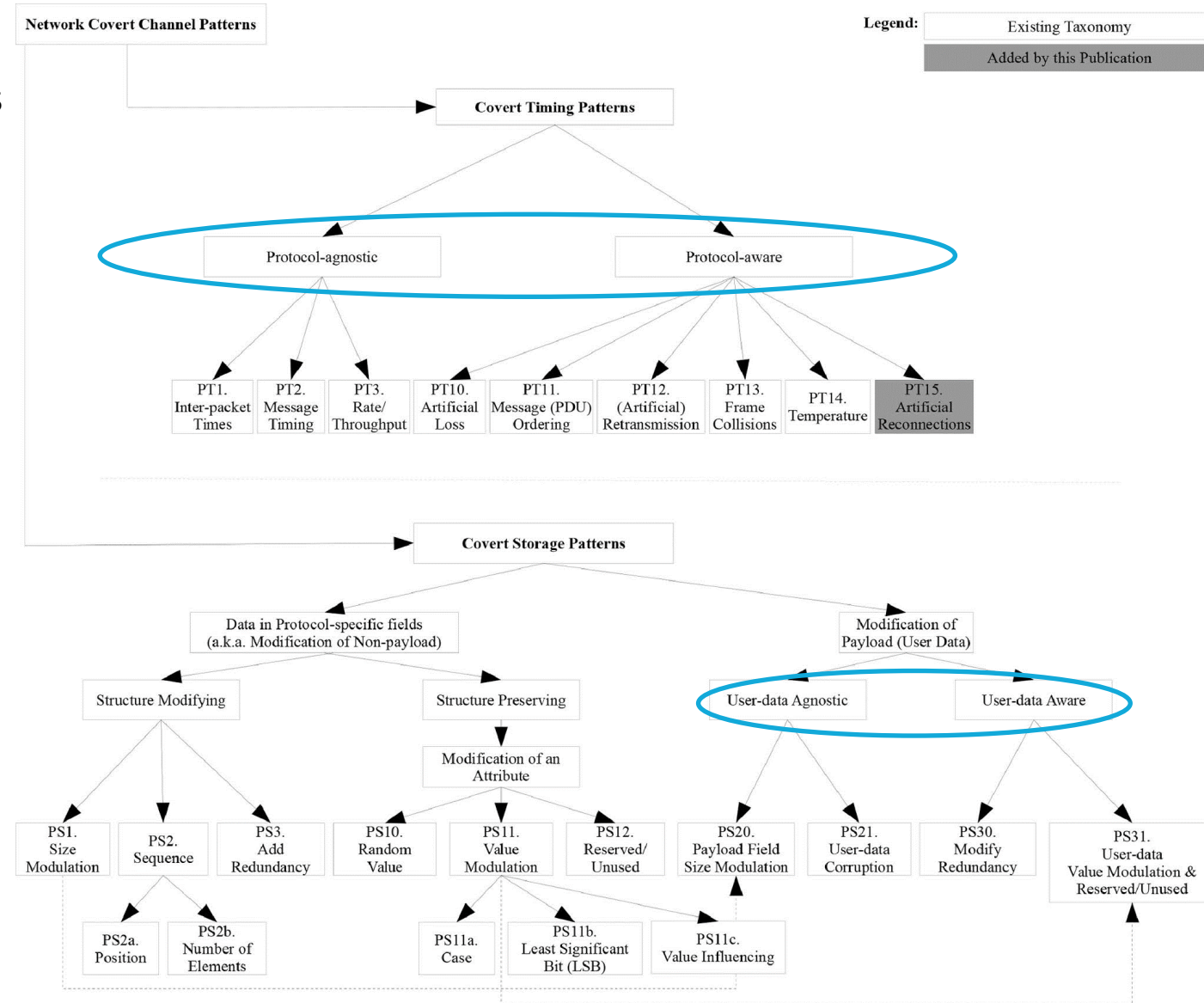
W. Mazurczyk, S. Wendzel, S. Zander et al.: [Information Hiding in Communication Networks](#), WILEY-IEEE, 2016, Chapter 3.

W. Mazurczyk, S. Wendzel, K. Cabaj: [Towards Deriving Insights into Data Hiding Methods Using Pattern-based Approach](#), in Proc. ARES, pp. 10:1-10:10, ACM, 2018.

A. Velinov, A. Mileva, S. Wendzel, W. Mazurczyk: [Covert Channels in MQTT-based Internet of Things](#), IEEE ACCESS, Vol. 7, pp. 161899-161915, 2019.

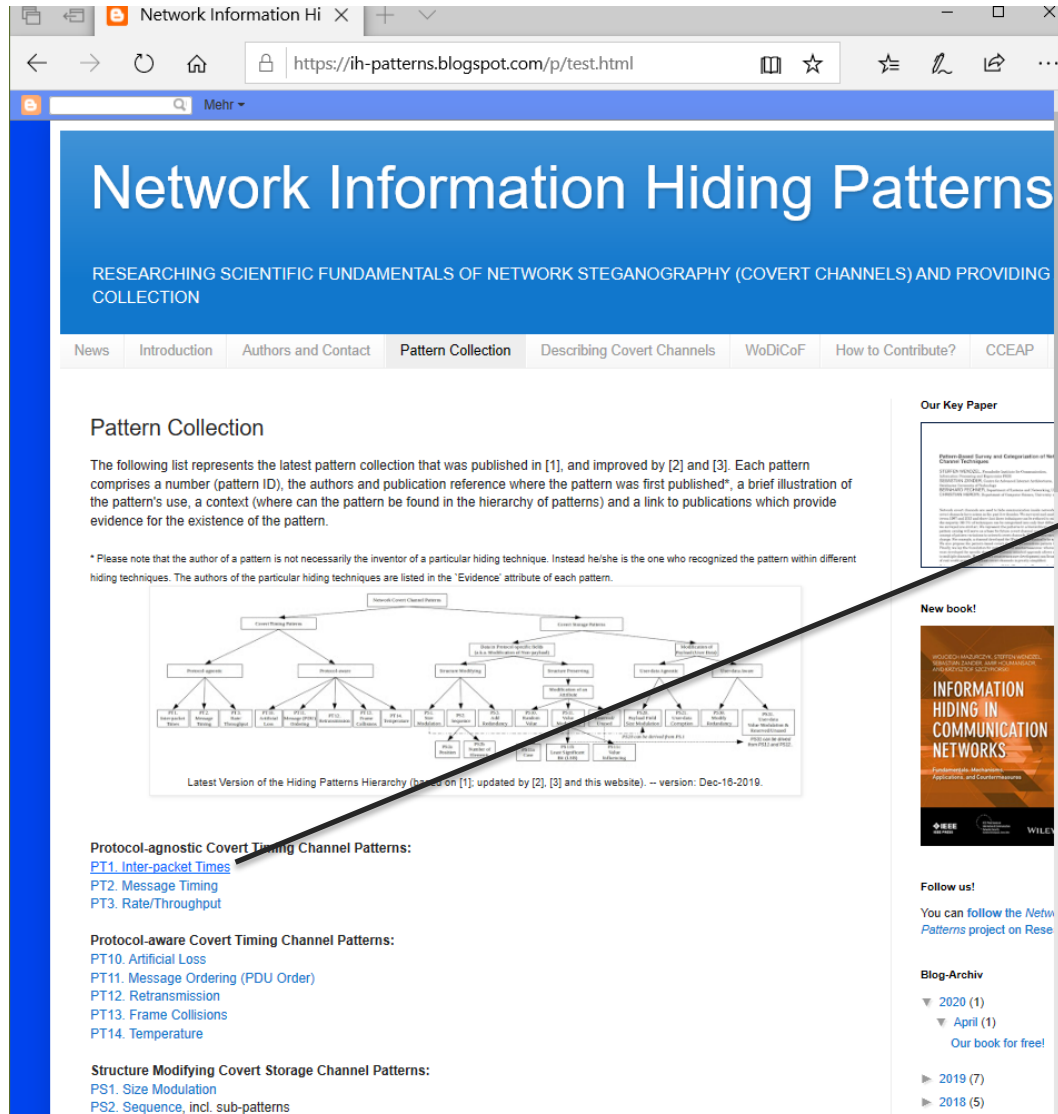
[Official Hiding Patterns Website \(https://ih-patterns.blogspot.com\)](https://ih-patterns.blogspot.com) ← always has the latest version of the taxonomy.

A. Mileva, A. Velinov, L. Hartmann et al. Comprehensive Analysis of MQTT 5.0 Susceptibility to Network Covert Channels. Computers & Security (COSE), Vol. 104, Elsevier, 2021.



Latest Version of Pattern Taxonomy

[Official Hiding Patterns Website \(https://ih-patterns.blogspot.com\)](https://ih-patterns.blogspot.com) ← always has the latest version of the taxonomy.



Network Information Hiding Patterns

RESEARCHING SCIENTIFIC FUNDAMENTALS OF NETWORK STEGANOGRAPHY (COVERT CHANNELS) AND PROVIDING COLLECTION

News Introduction Authors and Contact **Pattern Collection** Describing Covert Channels WoDiCoF How to Contribute? CCEAP

Pattern Collection

The following list represents the latest pattern collection that was published in [1], and improved by [2] and [3]. Each pattern comprises a number (pattern ID), the authors and publication reference where the pattern was first published*, a brief illustration of the pattern's use, a context (where can the pattern be found in the hierarchy of patterns) and a link to publications which provide evidence for the existence of the pattern.

* Please note that the author of a pattern is not necessarily the inventor of a particular hiding technique. Instead he/she is the one who recognized the pattern within different hiding techniques. The authors of the particular hiding techniques are listed in the 'Evidence' attribute of each pattern.

Protocol-agnostic Covert Timing Channel Patterns:

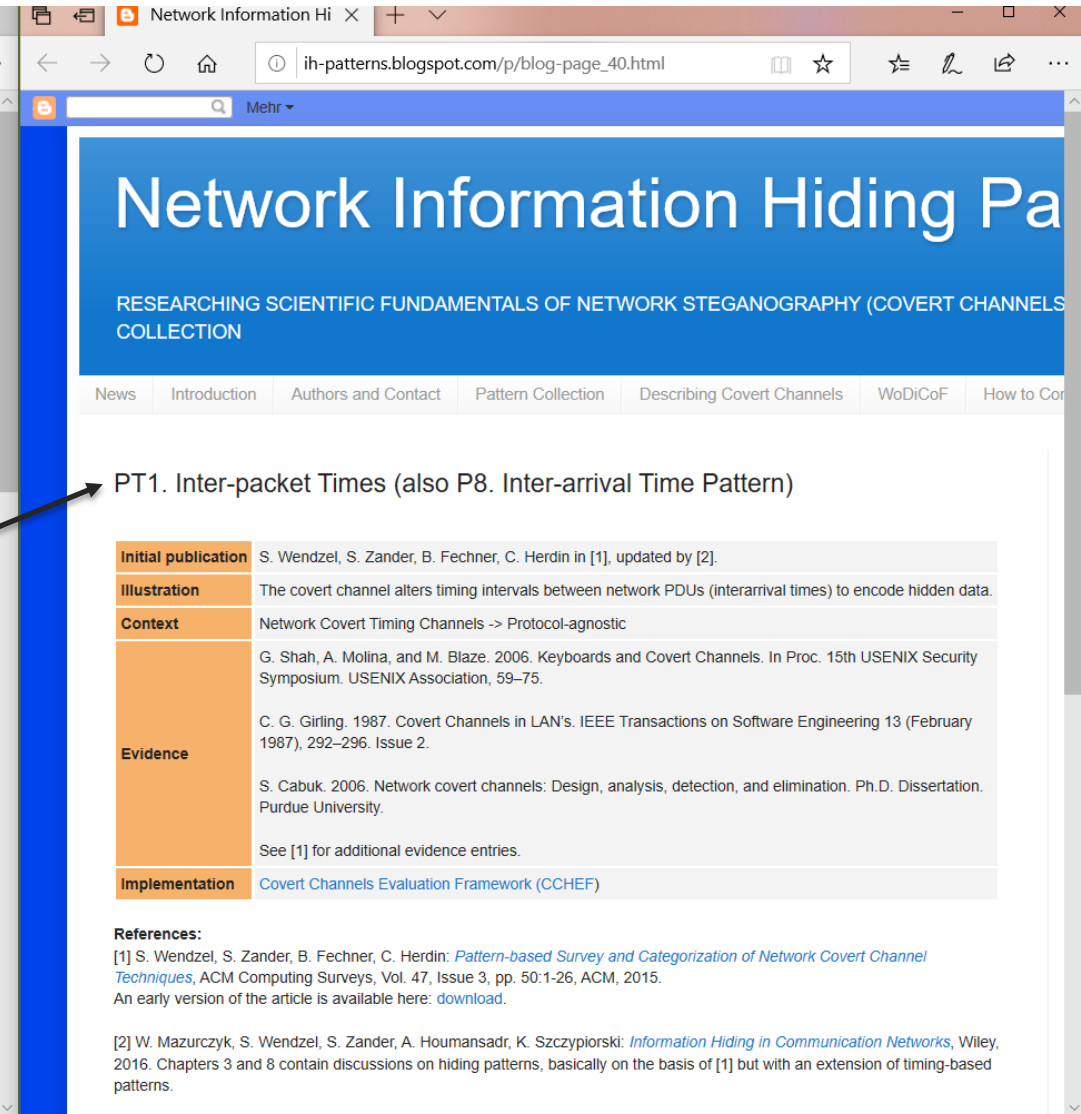
- [PT1. Inter-packet Times](#)
- [PT2. Message Timing](#)
- [PT3. Rate/Throughput](#)

Protocol-aware Covert Timing Channel Patterns:

- [PT10. Artificial Loss](#)
- [PT11. Message Ordering \(PDU Order\)](#)
- [PT12. Retransmission](#)
- [PT13. Frame Collisions](#)
- [PT14. Temperature](#)

Structure Modifying Covert Storage Channel Patterns:

- [PS1. Size Modulation](#)
- [PS2. Sequence, incl. sub-patterns](#)



Network Information Hiding Pa

RESEARCHING SCIENTIFIC FUNDAMENTALS OF NETWORK STEGANOGRAPHY (COVERT CHANNELS) COLLECTION

News Introduction Authors and Contact **Pattern Collection** Describing Covert Channels WoDiCoF How to Con

PT1. Inter-packet Times (also P8. Inter-arrival Time Pattern)

Initial publication	S. Wendzel, S. Zander, B. Fechner, C. Herdin in [1], updated by [2].
Illustration	The covert channel alters timing intervals between network PDUs (interarrival times) to encode hidden data.
Context	Network Covert Timing Channels -> Protocol-agnostic
Evidence	<p>G. Shah, A. Molina, and M. Blaze. 2006. Keyboards and Covert Channels. In Proc. 15th USENIX Security Symposium. USENIX Association, 59–75.</p> <p>C. G. Girling. 1987. Covert Channels in LAN's. IEEE Transactions on Software Engineering 13 (February 1987), 292–296. Issue 2.</p> <p>S. Cabuk. 2006. Network covert channels: Design, analysis, detection, and elimination. Ph.D. Dissertation. Purdue University.</p> <p>See [1] for additional evidence entries.</p>
Implementation	Covert Channels Evaluation Framework (CCEF)

References:

[1] S. Wendzel, S. Zander, B. Fechner, C. Herdin: *Pattern-based Survey and Categorization of Network Covert Channel Techniques*, ACM Computing Surveys, Vol. 47, Issue 3, pp. 50:1-26, ACM, 2015.
An early version of the article is available here: [download](#).

[2] W. Mazurczyk, S. Wendzel, S. Zander, A. Houmansadr, K. Szczypiorski: *Information Hiding in Communication Networks*, Wiley, 2016. Chapters 3 and 8 contain discussions on hiding patterns, basically on the basis of [1] but with an extension of timing-based patterns.

Limitations of the **CURRENT TAXONOMY**

Limitations of the Current Taxonomy

- **Focus limited to network communications**, neglecting other domains of steganography.
- Level of abstraction does **not allow for inclusion of *non-network* patterns**.
 - E.g., *user-data* from network perspective might be a digital media *payload*.
- Current taxonomy does not differentiate between the **embedding** and the **representation** process, rendering the interpretation of patterns ambiguous.
- If such a differentiation would be applied, some current patterns must be considered as **hybrid** patterns.

However, several of the current key concepts were kept; some patterns were renamed/made more abstract.

Improvements of Current Taxonomy During Years

- 2014/2015:
 - Steffen Wendzel, Sebastian Zander, Bernhard Fechner, Christian Herdin: **Pattern-based Survey and Categorization of Network Covert Channel Techniques**, Computing Surveys (CSUR), ACM, Vol. 47(3).
 - Definition of Hiding Patterns
 - First Taxonomy
 - Presents Methodology and Concepts
- 2016:
 - Wojciech Mazurczyk, Steffen Wendzel, Sebastian Zander, Amir Houmansadr, and Krzysztof Szczypiorski. **Information Hiding in Communication Networks: Fundamentals, Mechanisms, and Applications**. Chapter 3, Wiley-IEEE.
 - Several Improvements to the 2015-taxonomy
- 2018:
 - Wojciech Mazurczyk, Steffen Wendzel, and Krzysztof Cabaj. 2018. **Towards Deriving Insights into Data Hiding Methods Using Pattern-based Approach**. In Proc. Second International Workshop on Criminal Use of Information Hiding (CUING). ACM, 10:1–10:10.
 - New Patterns; New Categorizations
 - Hybrid Patterns
 - Distributed Hiding Patterns
- 2019:
 - Aleksandar Velinov, Aleksandra Mileva, Steffen Wendzel, Wojciech Mazurczyk: **Covert Channels in MQTT-based Internet of Things**, ACCESS, IEEE, Vol. 7.
 - New Sub-pattern
- 2020:
 - Mario Hildebrandt, Robert Altschaffel, Kevin Lamshöft, Mathias Lange, Martin Szemkus, Tom Neubert, Claus Vielhauer, Yongdian Ding, and Jana Dittmann. **Threat Analysis of Steganographic and Covert Communication in Nuclear I&C Systems**. In International Conference on Nuclear Security: Sustaining and Strengthening Efforts.
 - First Work on CPS Hiding Patterns
- 2021:
 - Aleksandra Mileva, Aleksandar Velinov, Laura Hartmann, Steffen Wendzel, and Wojciech Mazurczyk. 2021. **Comprehensive Analysis of MQTT 5.0 Susceptibility to Network Covert Channels**. Computers & Security (COSE), Vol. 104, Elsevier.
 - New Pattern (Artificial Retransmissions)

Improvements of Hiding Pattern-based Methodology

- 2015:
 - Steffen Wendzel, Carolin Palmer: **Creativity in Mind: Evaluating and Maintaining Advances in Network Steganographic Research**, J.UCS Vol. 21(12).
 - How to tell whether a new hiding technique represents a new pattern or solely uses an already existing pattern?
 - How patterns can be used during peer review.
- 2016:
 - Steffen Wendzel, Wojciech Mazurczyk, Sebastian Zander: **Unified Description Method for Network Information Hiding Methods**, J.UCS Vol. 22(11).
 - Making papers on hiding methods replicable and understandable (using patterns)
 - Steffen Wendzel, Wojciech Mazurczyk: **An Educational Network Protocol for Covert Channel Analysis Using Patterns (Poster)**, ACM CCS 2016.
 - How to teach in academia using hiding patterns.
- 2019:
 - Steffen Wendzel, Florian Link, Daniela Eller, Wojciech Mazurczyk: **Detection of Size Modulation Covert Channels Using Countermeasure Variation**, J.UCS, Vol. 25(11).
 - How to transfer a countermeasure that works for one pattern to another one?
- Excluded:
 - Several other papers addressing some pattern-related details.

APPROACH

1. Embedding and Representation Patterns



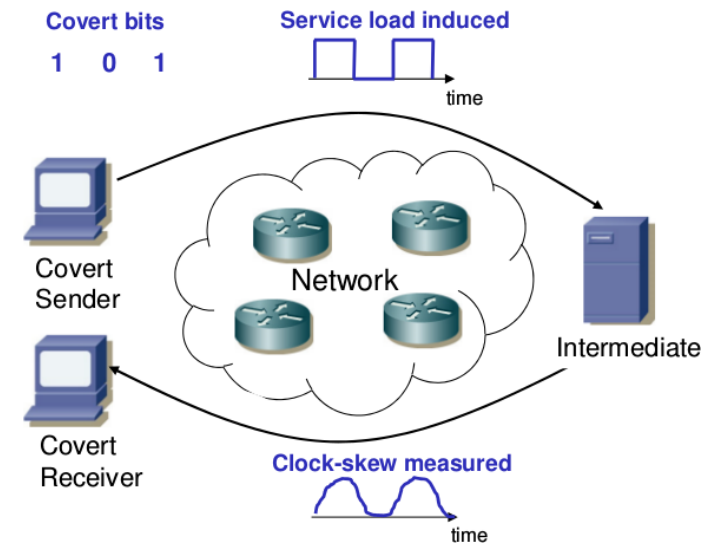
Embedding Patterns describe how secret information is embedded into a cover object, such as an image file or a network packet.



Representation Patterns describe how the secret information is represented in the cover object.

1. Embedding and Representation Patterns

- Why is the differentiation necessary?
- Sometimes, it isn't: Embedding Pattern = Representation Pattern
 - CS sends an IP packet to CR and manipulates the least significant bit of the TTL field. The embedding then follows the Value Modulation pattern while the data is also represented by the Value Modulation pattern.
- But sometimes, there is indeed a difference: Embedding Pattern \neq Representation Pattern
 - Assume an indirect CC as shown in the figure.
 - CS might influence the Intermediate node via the Rate/Throughput pattern.
 - CR monitors CPU load via different pattern, such as Inter-packet Times or Value Modulation.



2. Changing Pattern Names

- Several of the original patterns represent suitable concepts but need a naming and description that is not specific to the network context. We renamed several of them.
 - **Inter-packet Times** → **Event/Element Interval Modulation**
 - **Message Timing** → **Event Occurrence**
 - ...
- Some terms in the categorization had to be dropped as well, such as „payload“ and „protocol-awareness“.
- No differentiation between syntax and semantics.
 - Previously, one differentiated between structure-preservation and structure-modification. However, some patterns can modify both (see paper for an example).

2. Changing Pattern Names

- Finally, we came up with a glossary to make things more precise.
- Each pattern name is composed of an **identifier** and a **name**.
- **Identifier** has the Form **[ER][TN]n{D}**
 - **E**: Embedding Pattern **R**: Representation Pattern
 - **T**: Temporal Pattern **N**: Non-temporal Pattern
 - **n**: number of the pattern (can contain sub-numbers, e.g. **ET1.2.3**).
 - **D**: stego domain (only representation patterns; e.g. n=network, t=text stego, d=digital media, ...)
 - Example: **RT1t** = Representation Pattern, Temporal, Number 1, Text Steganography.

2. Changing Pattern Names

- **Name** contains the **identifier**, a modifiable **object** (e.g. *event* or *feature*) followed by an **action** (e.g. *modulation* or *occurrence*).
 - Example: **ET2. Event Occurrence**

(1) Modifiable Objects (see, Tab. 1):

- An *Event* describes a (timed or forced) appearance, which can be composed of several elements, e.g., 1) the appearance of a predefined character sequence; 2) a predefined specific sound in a video; 3) network connection establishment, reset or disconnection.
- An *Element* represents a single unit of a whole sequence, e.g., 1) a word/character of a text; 2) a pixel of an image; 3) a network packet of the whole flow.
- A *Feature* characterizes a property of an element to be modulated, e.g., 1) the color of a character; 2) the attribute of a tag in vector graphics; 3) the field / the size of a network packet.
- An *Interval* specifies the temporal gap between two events, e.g., 1) the duration of an audio file; 2) the time between sending a message and receiving the related acknowledgement.
- A *State/Value* denotes a non-temporal numerical or positional quantity of an element, feature, or event, e.g., 1) the values of TCP header fields (feature value); 2) the x-y-z coordinates of a player in a 3D game.

(2) Actions:

- An *Occurrence* is the temporal location of a given element, feature, or event observed in the cover.
- A *Modulation* of an element's (or event's) value (or state) is the selection of one particular value/state (out of multiple possible values/states).
- A *Corruption* refers to the blind overwriting of an element, feature or state/value.
- *Enumeration* means that the overall number of appearances of something is altered.
- *Repeating* refers to duplicating elements, events or features (multiple times). It can be considered a sub-form of the *enumeration* action.
- *Positioning* selects the non-temporal position of an element in a sequence of elements.

2. Changing Pattern Names

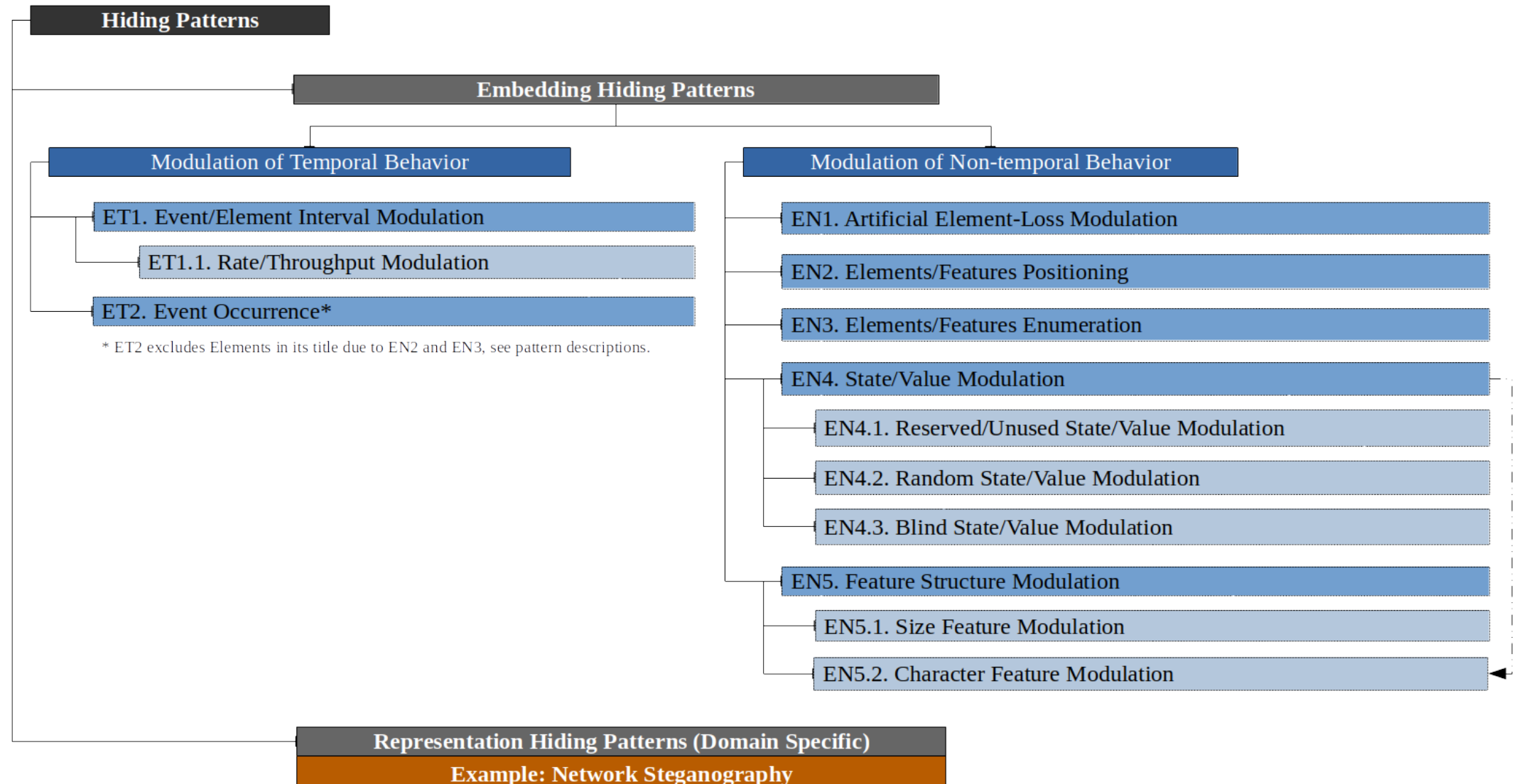
- **Name** contains the **identifier**, a modifiable **object** (e.g. *event* or *feature*) followed by an **action** (e.g. *modulation* or *occurrence*).
 - Example: **ET2. Event Occurrence**

Table 1: Differentiation between the types of *objects* used in this paper.

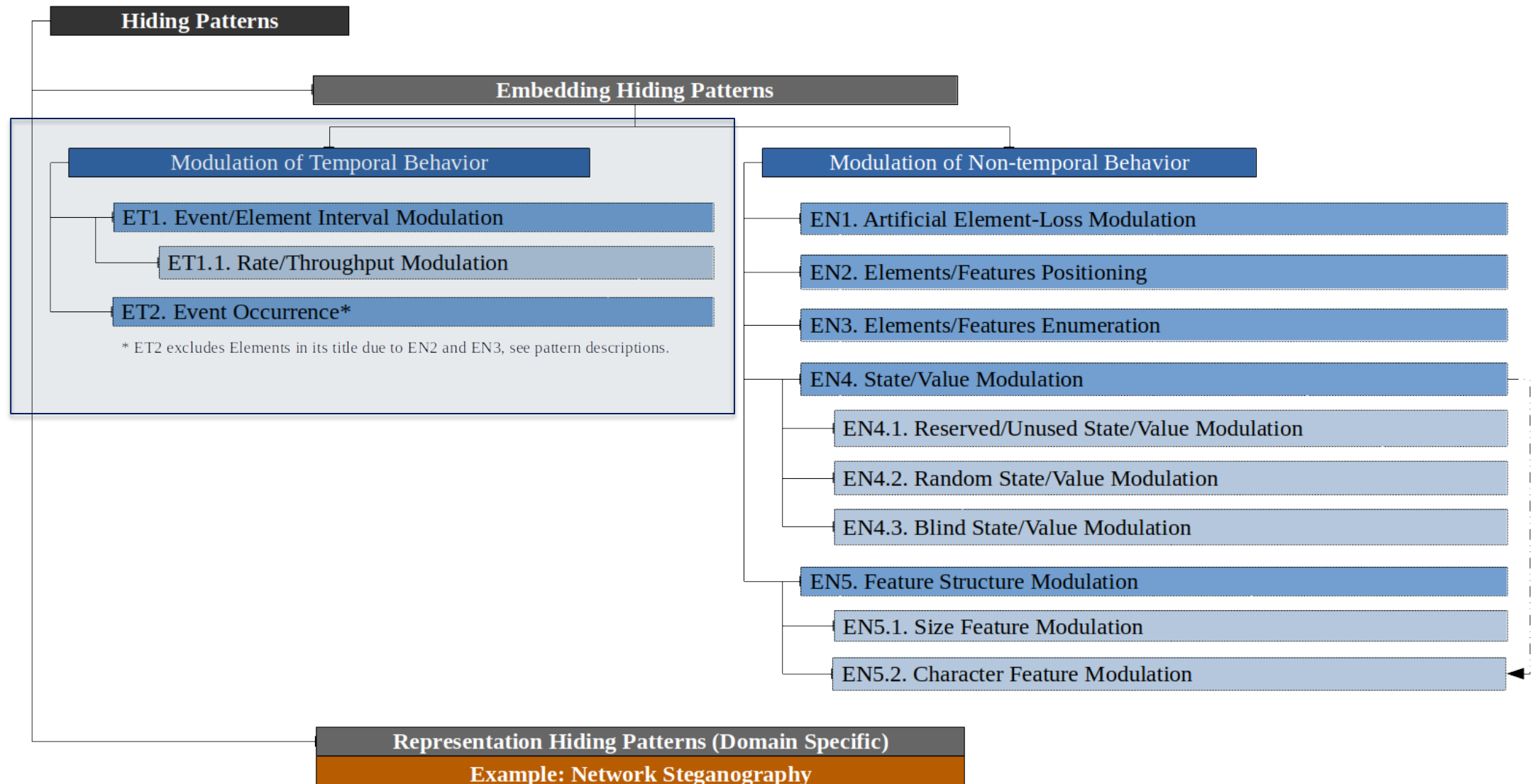
Domain	Interval	Event	Element	Feature	State/Value
network steganography	time between packets	presence of flow; disconnect	network packet	size of packet; field of packet	value of header field; number of packets
text steganography	time between text notes sent	occurrence of character sequence	character	color of character	number of characters
digital media steganography	duration of audio file	occurrence of pre-defined sound in MP3 file	pixel of image	color of pixel	value of pixel; number of pixels in image

Taxonomy of **EMBEDDING PATTERNS**

Embedding Patterns



Embedding Patterns

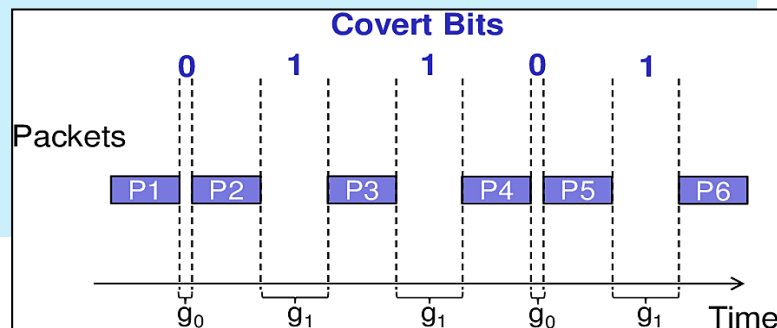


Temporal Hiding Patterns

ET1. Event/Element Interval Modulation

- The covert message is embedded by modulating the gaps between succeeding events/elements.
- Examples:
 1. modulating the inter-packet gap between succeeding network packets (elements) or between connection establishments (events);
 2. modulating the time-gap between succeeding cyber-physical actions, such as acoustic beeps.

Sub-patterns:
ET1.1 Rate/Throughput
Modulation

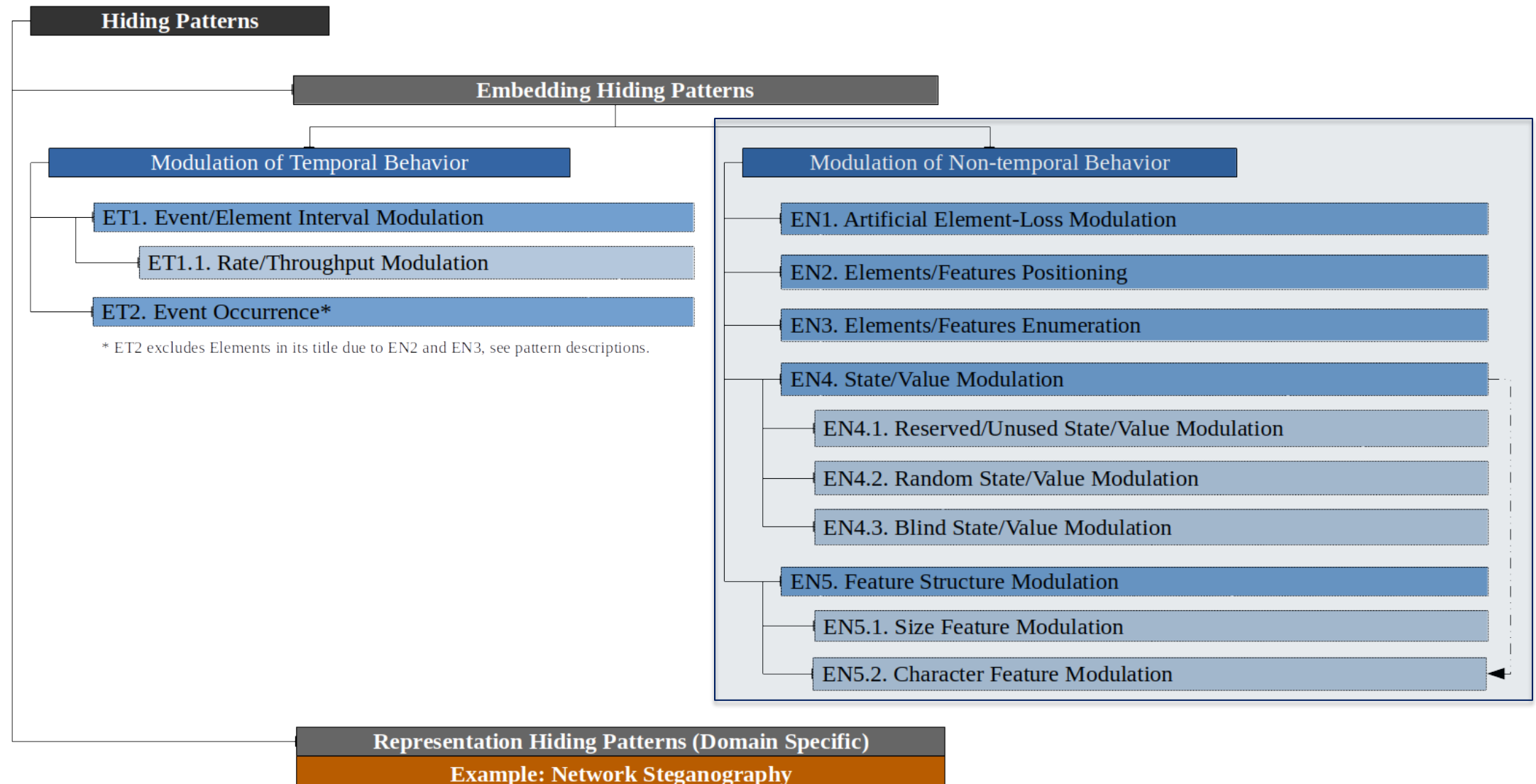


ET2. Event Occurrence

- The covert message is encoded in the temporal location of events (*in comparison to ET1.1, the rate of events is not directly modulated but events are triggered at specific moments in time, moreover, ET2 can be a single event while ET1.1 needs a sequence of elements*).
- Examples:
 1. sending a specific network packet at 6pm;
 2. influencing the time at which a drone starts its journey to some destination;
 3. performing a disconnect at a certain time.

Figures.: W. Mazurczyk, S. Wendzel, S. Zander et al.: Information Hiding in Communication Networks, Wiley-IEEE, 2016.

Non-temporal Embedding Patterns



Temporal Hiding Patterns

EN1. Artificial Element-Loss Modulation

- The covert message is embedded by modulating the artificial loss of elements.

- Examples:

- dropping TCP segments with an even sequence number;



- removing commas in sentences [1].

EN2. Elements/Features Positioning

- The covert message is embedded by modulating the position of a predefined (set of) element(s)/feature(s) in a sequence of elements/features.

- Examples:

- position of a HTTP header line in the list of optional header lines;
- placing a specific character in a paragraph.

```
GET HTTP/1.1
Host: mywebsite.xyz
User-Agent: MyBrowser/1.2.3 } s1
Accept-Language: en-US
```

```
GET HTTP/1.1
Host: mywebsite.xyz
Accept-Language: en-US } s2
User-Agent: MyBrowser/1.2.3
```

EN3. Elements/Features Enumeration

- The covert message is embedded by altering the overall number of appearances of elements or features in a sequence.
- Examples:
 1. fragmenting a network packet into either n or m ($n \neq m$) fragments;
 2. modulating the number of people wearing a t-shirt in a specific color in an image file;
 3. repeating an element/feature by duplicating a white space character (or not) in a text [1].

EN4. State/Value Modulation

- The covert message is embedded by modulating the states or values of features.
- Examples:
 1. modulating physical states, such as proximity, visibility, force, height, acceleration, speed, etc. of certain devices
 2. changing values of the network packet header fields (e.g., target IP address of ARP [2], Hop Count value in IPv6 [3] or the LSB in the IPv4 TTL);
 3. modulate the x-y-z coordinates of a player in a 3D multiplayer online game [4].

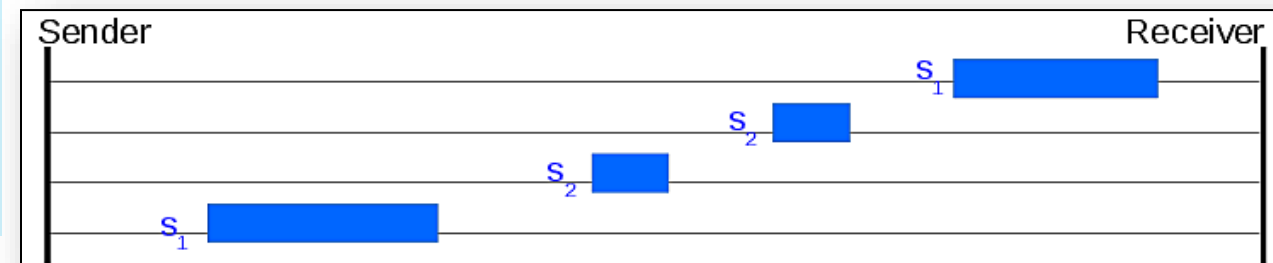
Pattern has three sub-patterns!

[1] Walter Bender, Daniel Gruhl, Norishige Morimoto, and Anthony Lu. 1996. Techniques for data hiding. IBM Systems Journal 35 (Nos3&4) (1996). ||| [2] Liping Ji, Yu Fan, and Chuan Ma. 2010. Covert channel for local area network. In 2010 IEEE International Conference on Wireless Communications, Networking and Information Security. ||| [3] Norka B Lucena, Grzegorz Lewandowski, and Steve J Chapin. 2005. Covert channels in IPv6. In International Workshop on Privacy Enhancing Technologies. Springer, ||| [4] Sebastian Zander, Grenville Armitage, and Philip Branch. 2008. Covert channels in multiplayer first person shooter online games. In 2008 33rd IEEE Conference on Local Computer Networks (LCN). IEEE.

EN5. Feature Structure Modulation

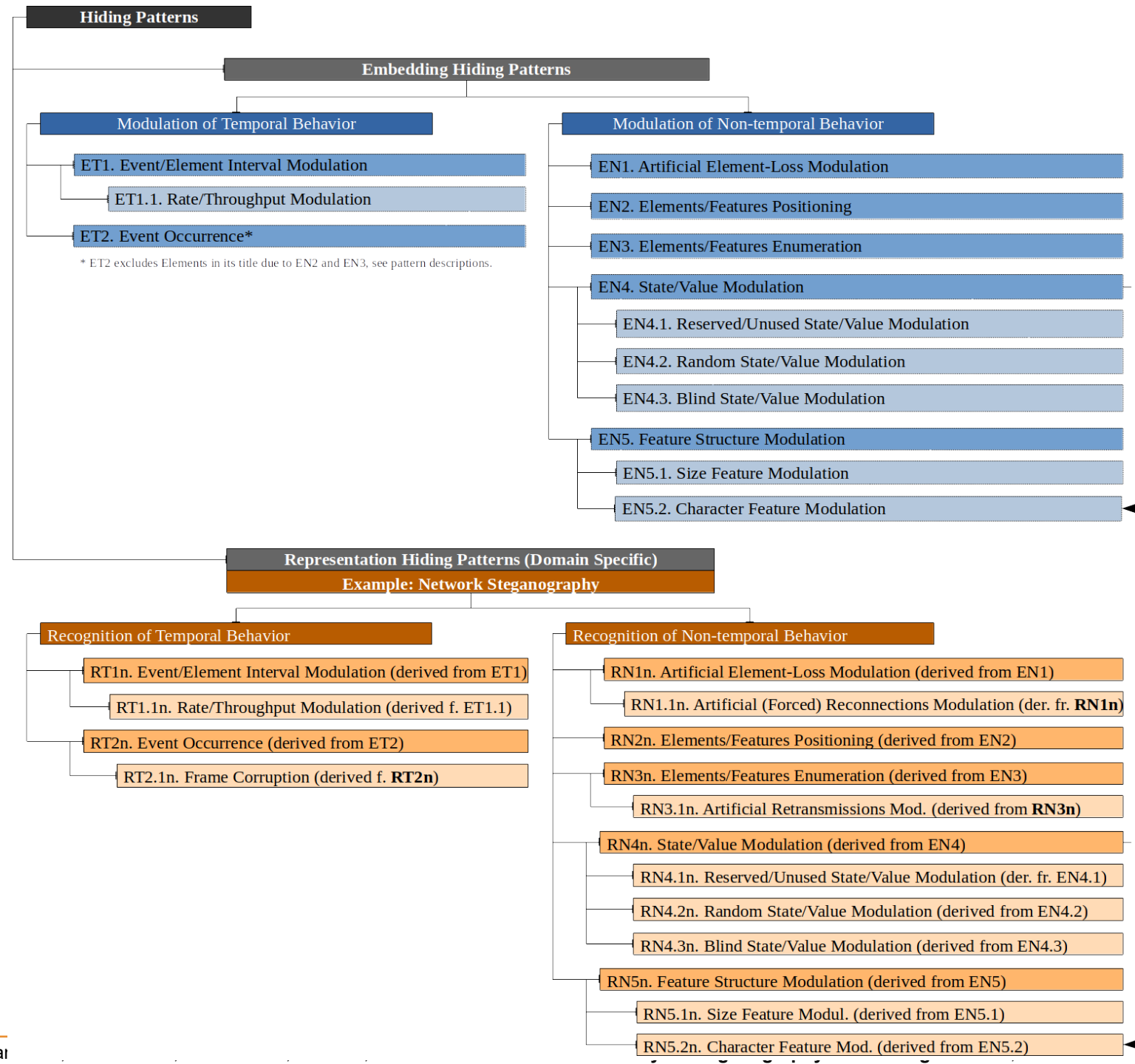
- This hiding pattern comprises all hiding techniques that modulate the structural properties of a feature (but not states/values (EN4), positions (EN2) or number of appearances (EN3)).
- Examples:
 1. increasing/decreasing the size of succeeding network packets;
 2. changing the color/style of characters in texts.

Pattern has two sub-patterns.

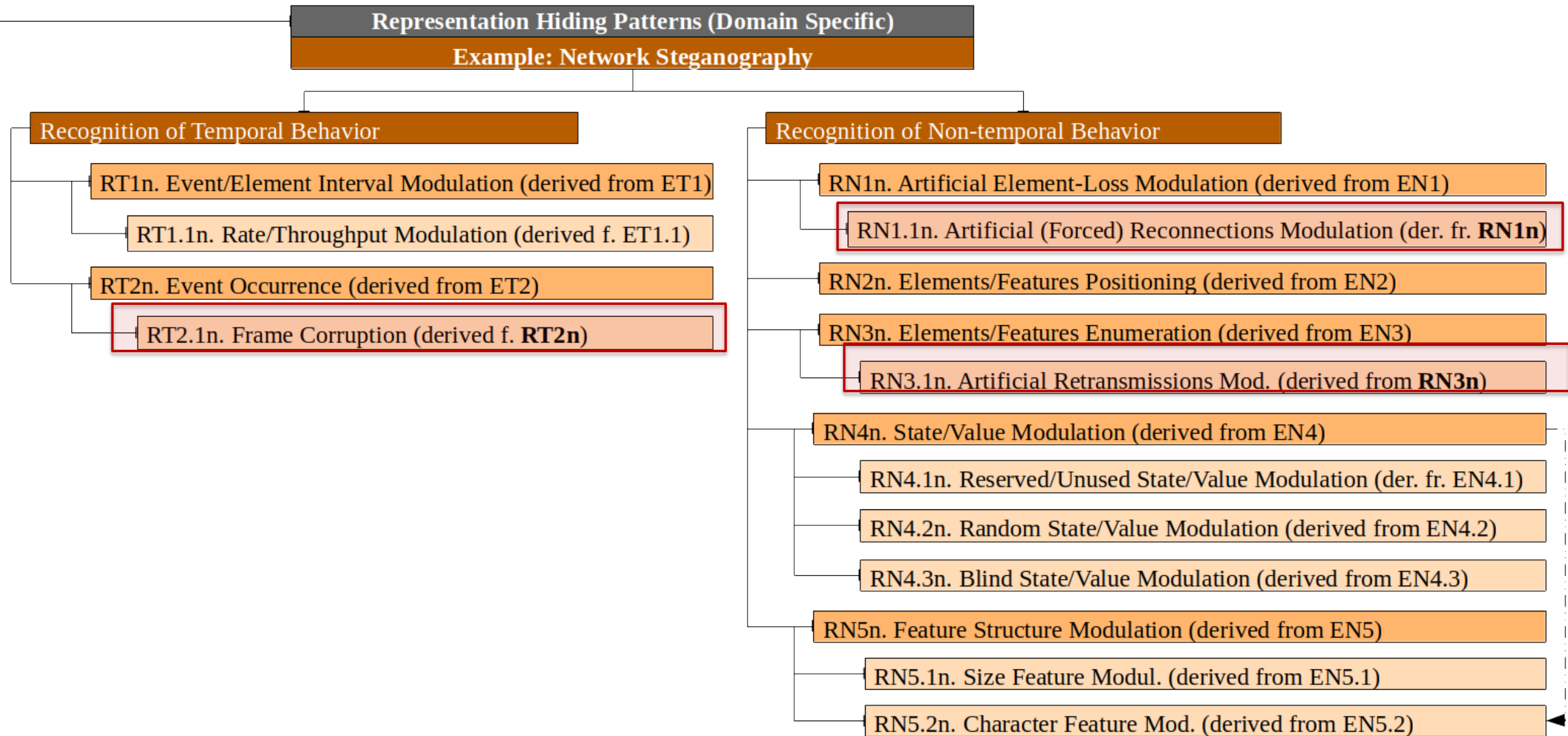


Representation Patterns Example Using **NETWORK STEGANOGRAPHY**

Example of Representation Patterns: Network Steganography



Let us have a look at the differences:



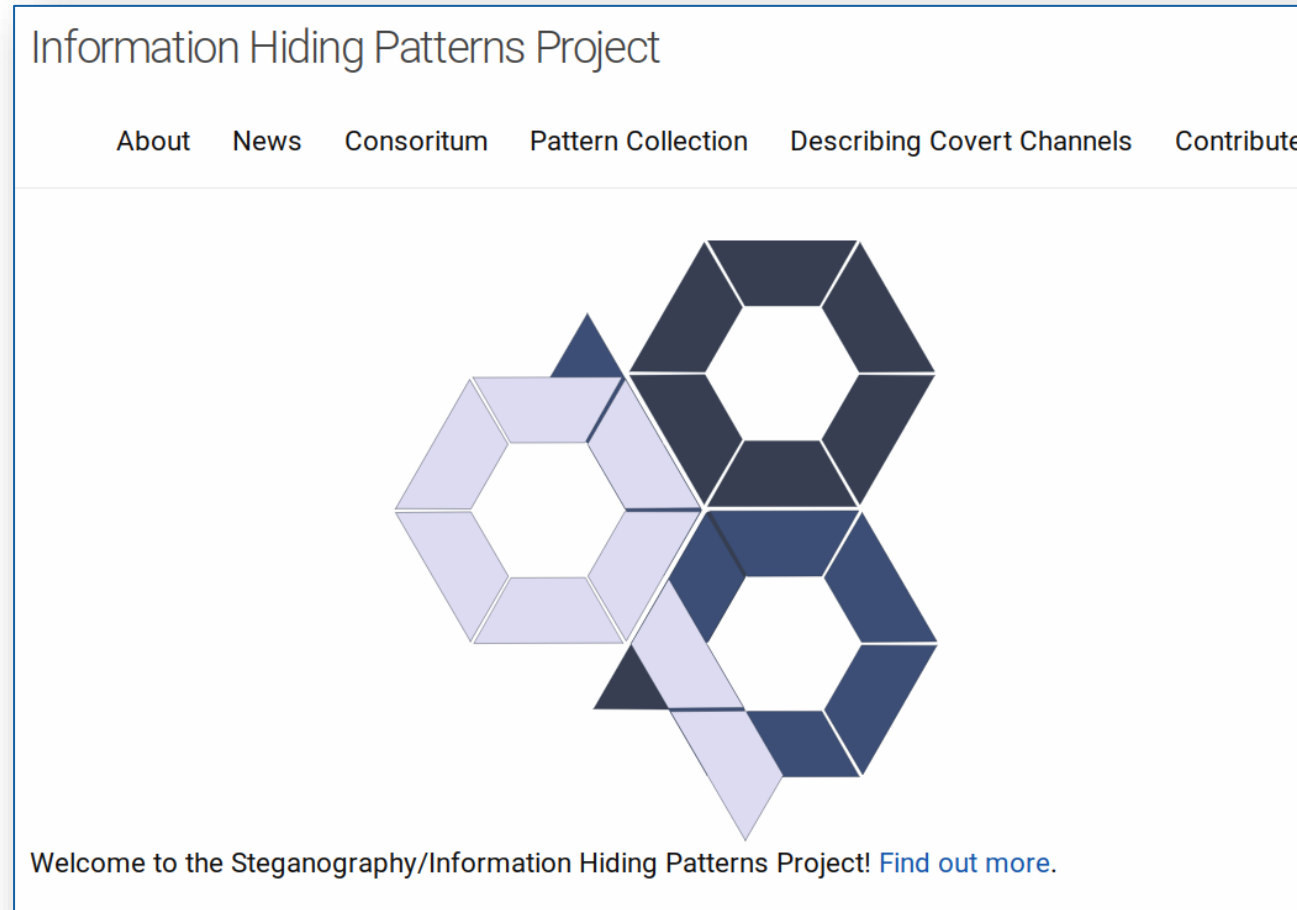
Conclusions & **FUTURE WORK**

Conclusions & Future Work

- **Revised existing taxonomy** of hiding patterns
- **Addressed several limitations** of the current taxonomy
- Made the **taxonomy more general** so that it **can be applied by all steganography domains**, instead of solely network steganography.
- **Kept the good concepts** and pattern-ideas **of the existing taxonomy** whenever feasible (do not re-invent the wheel).
- First taxonomy of representation patterns for **network steganography**.
- Future Work:
 - Add representation patterns for **digital media, text, filesystem and CPS steganography** (first discussions on digital media and text stego can be found in the paper).
 - Potentially extend size of consortium to see whether additional information hiding domains, such as **digital watermarking**, can be integrated.

Thank you for your kind attention!

<https://patterns.ztt.hs-worms.de>



Acknowledgements

- Parts of the work from Brandenburg and Magdeburg authors in this paper (i.e., on definitions and general discussions) have been funded by the German Federal Ministry for Economic Affairs and Energy (**BMWi, Stealth-Szenarien, Grant No. 1501589A and 1501589C**) within the scope of the German Reactor-Safety-Research-Program.
- Parts of the work of Laura Hartmann has been funded by the **European Union from the European Regional Development Fund (EFRE)** and the **State of Rhine-land-Palatinate (MWWK), Germany. Funding content: P1-SZ2-7 F&E: Wissens- und Technologietransfer (WTT), Application number: 84003751, project MADISA.** Her work has also been funded by **Programm zur Förderung des Forschungspersonals, Infrastruktur und forschendem Lernen (ProFIL) of the University of Applied Sciences Worms.**
- Parts of the work of Luca Caviglione and Wojciech Mazurczyk have been supported by the **SIMARGL Project - Secure Intelligent Methods for Advanced RecoGnition of malware and stegomalware**, with the support of the European Commission and the **Horizon 2020 Program**, under **Grant Agreement No. 833042.**