(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 December 2003 (24.12.2003)

PCT

(10) International Publication Number
WO 03/107156 A2

(51) International Patent Classification[7]: G06F 1/00, H04L 29/06

(21) International Application Number: PCT/US03/19217

(22) International Filing Date: 17 June 2003 (17.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/390,683 18 June 2002 (18.06.2002) US

(71) Applicant (*for all designated States except US*): **HONEYWELL INTERNATIONAL INC.** [US/US]; 101 Columbia Road, Morristown, NJ 07962 (US).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): **PHINNEY, Thomas, L.** [US/US]; 5012 West Torrey Pines Circle, Glendale, AZ 85308 (US).

(74) Agent: **MIOLOGOS, Anthony**; Honeywell International Inc., 101 Columbia Road, Morristown, NJ 07962 (US).

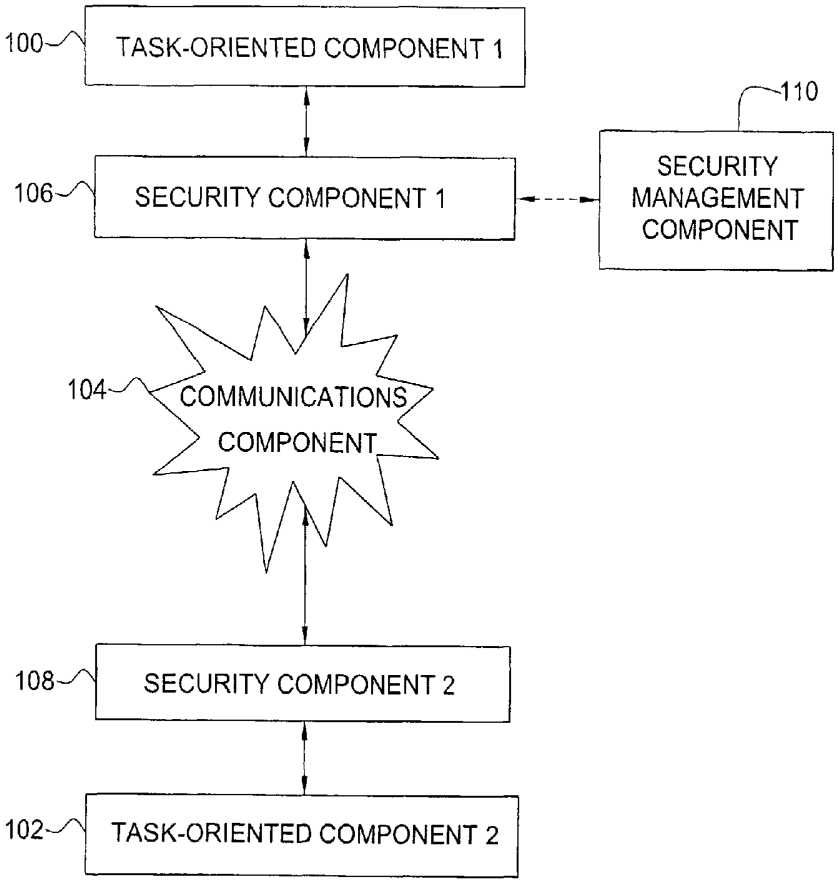
(81) Designated States (*national*): AT, CA, FI, JP, KR, NO, US.

(84) Designated States (*regional*): Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR CONFIGURING AND COMMISSIONING CSMs



(57) Abstract: A method for loading cryptographic protocols and installing a ComSec master dongle (CSM) to provide secure communications in a control system, such as a supervisory control and data acquisition (SCADA) system with a wide area network (WAN) is disclosed.

METHOD FOR CONFIGURING AND COMMISSIONING CSMs

CROSS REFERENCE

This application claims priority of U.S. Provisional Patent Application Serial No. 60/390683, filed on June 18, 2002, entitled "METHOD FOR SCADA COMSEC," which is incorporated herein by reference.

This application is further related to co-pending and co-owned patent applications entitled: "SYSTEM AND METHOD FOR SECURING NETWORK COMMUNICATIONS," Honeywell Docket No. H18-03434, U.S. Serial No. 10/___,___; "MASTER DONGLE FOR A SECURED DATA COMMUNICATIONS NETWORK," Honeywell Docket No. I20-04611, U.S. Serial No. 10/___,___; "DONGLE FOR A SECURED DATA COMMUNICATIONS NETWORK," Honeywell Docket No. I20-04612, U.S. Serial No. 10/___,___; "METHOD FOR CONFIGURING AND COMMISSIONING CSSs," Honeywell Docket I20-04614, U.S. Serial No. 10/___,___; "METHOD FOR ESTABLISHING SECURE NETWORK COMMUNICATIONS," Honeywell Docket No. I20-04615, U.S. Serial No. 10/___,___, all filed on June 17, 2003, and all having a common assignee as the present invention.

BACKGROUND

1. Field of the Invention

The present invention generally relates to communications security and relates in particular to configuring and commissioning ComSec master (CSM) devices.

2. Description of the Related Art

In an age of growing computer literacy and organized social disorder, there is an increasing need to protect corporate resources and national critical infrastructure from cyberattacks. For example, the electric power industry needs protection for the information carried on communication links between centralized control centers and outlying equipment sites.

Without such protection, an eavesdropping competitor, through modeling (for instance, with a neural network), can evaluate the rough economics of a system's operation and then use that knowledge of incremental cost to provide a bidding edge in the real-time marketplace. If eavesdropping is ongoing, this information advantage is magnified.

Without information protection, those of ill intent can determine the state of a system to select the most opportune moment and method of attack. More active assailants can take control of the communications and through it take control of the outlying sites. Through misrepresentation of the state of those outlying sites, they may also induce actions by the central control system and its operators that degrade or damage other parts of the system's operation or even its physical integrity.

There is an urgent need for cyber protection of such communication links, including:

1. Protecting communicated information from disclosure to unauthorized eavesdroppers;
2. Detecting and rejecting messages that originated from an unauthorized source or were altered in transit by an unauthorized source; and
3. Detecting and rejecting unaltered messages that originated from an authorized source when they were recorded but then replayed at a later time.

Any system that protects electronic communications against unauthorized message senders needs to be fail-safe so that unauthorized messaging is still rejected after potential failure conditions. Otherwise, an organized attacking group can take over field sites simply by intercepting the transmission paths, such as a telephone switching site or microwave relay, and substituting its own messages.

The ability to initiate such an attack can be put in place and go undetected for months or years before any use. Telephone switches can and have been hacked. Trojaned equipment can be substituted for the original. In this modern era of multinational terrorists and state-funded cyberwarriors, such modes of attack cannot be discounted.

Once a threat is appreciated, however vaguely, protective measures can be planned and risks mitigated. New systems can be designed to reduce the threat. Cyberprotection for communications can be included in new designs from the start, provided the industry can agree on an adequate common approach for its multiple vendors to follow. Existing systems pose a different problem. In general, they cannot be redesigned and so must instead be retrofitted to protect against the threat. Therein lies the most difficult problem.

Even within a single company, the communication links that need to be protected typically use a heterogeneous collection of incompatible protocols implemented in multiple generations of equipment from a variety of vendors. The problem is further complicated with intertie of originally disjoint systems resulting from corporate mergers, asset transfers, and restructuring, as well as that resulting from centralizing control and maintenance for improved productivity.

Most of the existing communications equipment is itself too old to modify. In many cases, the designers are dead or long retired and sometimes the vendor companies themselves no longer exist. Standard industry practice is to use the existing equipment as long as possible, because there is little or no economic justification for replacing the old equipment. Any approach to providing cybersecurity for such equipment needs to address these constraints.

When additional equipment is inserted inline on a communications path, it imposes both physical and performance burdens on the system. The physical burdens are those of housing, powering, connecting, and maintaining the new equipment. The performance burdens are those caused by the delay in communications induced by the new equipment and by the unavoidable increase in the failure rate of the communications path.

The physical burden imposed by new equipment is a major concern. If it takes a crane or forklift operator to deliver an industrially-hardened enclosure, facilities personnel to install it, a communications technician to install the new equipment in the enclosure and to wire it into the existing communications system, and a licensed electrician to provide the equipment's power, the economic burden of adding cyberprotection is great.

Millions of systems in corporate resources and national infrastructure are vulnerable and unsecured. There is a critical need for a simple, fast, and economical method to protect both existing and new systems. Furthermore, there is a need for a security system that can be flexibly implemented in either hardware or software depending on the characteristics of the system being protected.

install the new equipment in the enclosure and to wire it into the existing communications system, and a licensed electrician to provide the equipment's power, the economic burden of adding cyberprotection is great.

Millions of systems in corporate resources and national infrastructure are vulnerable and unsecured. There is a critical need for a simple, fast, and economical method to protect both existing and new systems. Furthermore, there is a need for a security system that can be flexibly implemented in either hardware or software depending on the characteristics of the system being protected.

SUMMARY OF THE INVENTION

There is a method for configuring and commissioning. A communications protocol is specified for a first security component. An authentication method is specified for the first security component. Key escrow parameters are specified for the first security component. The first security component is a first ComSec master (CSM). The first security component is authorized to activate a commissioning method. While authorization is in force, a second security component is coupled to a port of the first security component and configured by the first security component.

In configuring the second security component by the first security component, a birth key encryption key (KEK) is requested from the second security component. The birth KEK is decrypted to establish a session key. The first security component generates an identifier for the second security component. The first security component generates a personal KEK. The first security component sends the identifier and the personal KEK to the second security component using the session key. The identifier is a unique system device identifier.

The first security component is sometimes a replacement security component. The replacement security component is connected to a network so that other security components share a session key with it. Sometimes, when the replacement security component is a replacement CSM, the second security component is a ComSec slave (CSS), other security components are other CSSs, and a CSM database is either not operational or not retrievable.

Upon powerup, the replacement CSM stops communication from a master terminal unit (MTU) and communication from the CSS to the replacement CSM. The replacement CSM requests an address from a remote terminal unit (RTU). The replacement CSM deciphers the address to determine whether the CSS is new. The list is associated RTU addresses, including any multicast addresses with associated session keys. The address is a unicast address for an enciphered version of a backup KEK for the CSS.

If the CSS is newly discovered by the replacement CSM, the CSM generates a new KEK for the new CSS, enciphering the new KEK under the personal KEK, and sending the new KEK to the new CSS. The replacement CSM receives at least one list(s) in at least one message protected by the new KEK from the CSS. The replacement CSM uses the list(s) received from the discovered CSS to regenerate at least a portion of the CSM database. The replacement CSM generates new session keys based on the list and sends the new session keys to all the CSSs such that each CSS receives only the session keys needed for communications with its associated remote terminal unit(s) (RTUs).

There is a method for deploying. A security component is installed on a connection of a task-oriented device to a modem by interrupting the connection to insert the security component between the task-oriented device and the modem. Power is applied to the security component and the security component alters a communication to or from the task-oriented device. The security component is configured and commissioned before it is installed. The task-oriented device is an MTU and the security component is a CSM.

These and other features, aspects, and advantages of the present invention will become better understood with reference to the following drawings, description, and appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of one embodiment of a system for securing network communications according to the present invention.

FIG. 2 is a block diagram of another embodiment of a system for securing network communications according to the present invention.

FIG. 3 is a block diagram of a preferred embodiment of a system for securing network communications according to the present invention.

FIG. 4 is a block diagram of another example of a system for securing network communications according to the present invention.

FIG. 5 is a block diagram of a method for configuring and commissioning according to the present invention.

FIG. 6 is a block diagram of a method for deploying according to the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the following detailed description, reference is made to the accompanying drawings. These drawings form a part of this specification and show by way of example specific preferred embodiments in which the present invention may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the present invention. Other embodiments may be used. Structural, logical, and electrical changes may be made without departing from the spirit and scope of the present invention. Therefore, the following detailed description is not to be taken in a limiting sense and the scope of the present invention is defined only by the appended claims.

FIG. 1 shows one embodiment of a system for securing network communications. Security is defined as measures taken to protect a system. Also, security is a condition that results from the establishment and maintenance of protective measures that ensure a state of inviolability from hostile acts or influences. In practical terms, security hinges on good encryption, but good encryption is by far not enough to obtain good security and a poorly-engineered system does not obtain sufficient security even though high-quality encryption might be employed. In addition, security is the condition of system resources being free from unauthorized access and from unauthorized or accidental change, destruction, or loss. In summary, security is the condition of a system that results from the establishment and maintenance of measures to protect the system.

In FIG.1, a first task-oriented component 100 and a second task-oriented component 102 have secure communications over a communications component 104, such as a network. The secure communications are enabled by a first security component 106 and a second security component 108 with the help of a security management component

110. First task-oriented component 100 and second task-oriented component 102 are any two pieces of equipment capable of communicating over a network, such as two computers. They are task-oriented in that they primarily perform some task unrelated to communications, such as process control or automation. Communications component 104 is any kind of symmetric or asymmetric communications system. Some examples are a local area network (LAN), a wide area network (WAN), and the like.

First security component 106 and second security component 108 may be implemented in either hardware, as a dongle, or in software and operate to alter a communication between first task-oriented component 100 and second task-oriented component 102 in order to secure the communication. A dongle is a device that is capable of being attached to a standard connector on a computer, a modem, or a similar piece of equipment. The dongle is sometimes a small, hard-shelled device. The dongle is typically interposed between the connector and any cable for other equipment that might normally be attached to that connector.

A communication from first task-oriented component 100 to second task-oriented component 102 is processed by first security component 106 to alter the communication in a certain way before it passes to communications component 104. Then, second security component 108 alters the communication from communications component 104 in such a way as to restore the communication back to its unaltered form. The communication is then passed to second task-oriented component 102. In this way, the alteration is transparent to the task-oriented components.

In some embodiments, first security component 106 is a communications security master (CSM) and second security component 108 is a communications security slave (CSS).

A ComSec master (CSM) is software and related hardware in a ComSec dongle master (CSM), or equivalent software and related hardware in a control system, such as a supervisory control and data acquisition (SCADA) master computer or controller. SCADA is a type of loosely-coupled distributed monitoring and control system commonly associated with electric power transmission and distribution systems, oil and gas pipelines, water and sewage systems, and other systems. A CSM performs several functions. First, a CSM configures and commissions each ComSec dongle slave (CSS) before deployment. Second, a CSM provides source authentication, confidentiality, integrity protection, and replay protection to the communications sent to and received from the deployed RTUs. Third, a CSM provides key management services, including key generation and key escrow, for the communications system. Fourth, a CSM provides code management services, including providing initial CSS code for non-dongle CSSs and code updates for all CSSs and other CSMs in the system. Finally, a CSM provides remote management, logging, and alarming of significant security events, via a network interface.

Authentication, confidentiality, integrity protection, and replay protection are various kinds of security. Authentication is any security measure designed to establish the validity of a transmission, message, or originator; also a means of verifying an individual's eligibility to receive specific categories of information. Confidentiality is the nonoccurrence of the unauthorized disclosure of information. Data integrity is the condition that exists when data is unchanged from its source and has not been accidentally or maliciously modified, altered, or destroyed. Data integrity protection is the degree to which a system or component detects unauthorized access to, or modification of, computer programs or data. Replay protection is validating message sequencing and timeliness so that prior valid messages cannot be replayed without detection of their lack of timeliness. A nonce is a random or non-repeating value that is included in data exchanged by a protocol, usually for the purpose of guaranteeing liveness and, thus, detecting and protecting against replay attacks. Spoofing is pretending to be another, as in one agent masquerading as another. More technically, spoofing is interception, alteration, and retransmission of a signal or data in such a way as to mislead the recipient.

A ComSec slave (CSS) is software and related hardware in a ComSec dongle for a remote terminal unit (RTU) or equivalent embedded software and assigned hardware in an RTU. A CSS provides source authentication, confidentiality, integrity protection, and replay protection to the communications received from and sent to the master terminal units (MTUs). A master terminal unit (MTU) is a master station in a control system. A remote terminal unit (RTU) is a remote station in a control system. In some embodiments, the CSM performs some or all of the functions of security management component 110.

Deploying is the act of taking a previously configured and commissioned CSS to the field, momentarily disconnecting a slave modem from its associated RTU(s), interposing the CSS dongle between the slave modem and the RTU(s), and reconnecting them all so that the RTU(s) are connected transitively through the CSS dongle to the modem. CSMs are similarly deployed.

Configuring is the act of writing the non-volatile memory of a CSS with the current revision of the CSS software appropriate for the communications protocol of the network.

Security management component 110 operates to manage first security component 106 and second security component 108 by managing recovery keys and acting as an originating key server and a code server. Security management component 110 has access to a random number generator, which is sometimes used to generate unpredictable encryption keys. In one embodiment, the security management component 110 is implemented as a key management center (KMC) in a computer that is physically secure, such as in a secured facility. A key management center (KMC) is a secured dedicated computer system connected to a network, such as the Internet for license authentication, initial secret key administration, and key recovery by a control system operator. A control system operator is a business enterprise responsible for operating a control system. The KMC is used to detect piracy and enforce licensing and to provide a service opportunity for a last-ditch remote dongle management reclamation service as well as to function as a key server and code server. The latter function is for code upgrades and to support new types of CSMs and CSSs. The dotted line connecting security management component 110 to security component 106 indicates that this communication is occasional rather than continuous.

A key is information (usually a sequence of random or pseudo-random binary digits) used initially to set up and periodically change the operations performed in cryptographic equipment or software for the purpose of encrypting or decrypting electronic signals. Key management is the process by which a key is generated, stored, protected, transferred, loaded, used, and destroyed. A secret key is the protected secret of secret key cryptography, used for both encryption and decryption. Secret key cryptography is a type of cryptography in which a shared secret is used for both encryption and decryption, in contrast with public key cryptography where different keys are used for encryption than for decryption.

FIG. 2 shows another embodiment of a system for securing network communications. In comparing FIG.1 and FIG. 2, in FIG. 2, the security components 106, 108 are inside task-oriented components 100 and 102 instead of being interposed between task-oriented components 100 and 102 and communications component 104, as in FIG. 1. For example, if first security component 106 is implemented in software and first task-oriented component 100 is a computer, then first security component 106 comprises executable instructions, keys, and key-related data stored in memory on the computer.

FIG. 3 shows a preferred embodiment of a system for securing network communications applied to a SCADA system. Like FIG. 1, FIG. 3 shows task-oriented components having secure communications over communications components. However, there are more task-oriented components and communications components in various configurations.

The general elements shown in FIG. 1 can be mapped onto the specific elements in FIG. 3. An example of first task-oriented component 100 of FIG. 1 is an MTU, such as MTU 300. An example of second task-oriented component 102 of FIG. 1 is an RTU, such as RTU 302. An example of communications component 104 of FIG. 1 is a plurality of networks and modems, such as network 304 and modems 305 and 307.

An example of security management component 110 of FIG. 1 is a KMC, such as a remote security management component KMC 310 coupled with a local security management component LKMC 311. The dotted line connecting KMC 310 to LKMC 311 indicates that this communication connection is occasional rather than continuous. The key server and code server functions are distributed so that, while they originate in the KMC 310, they are operationally either part of each CSM or part of a LKMC 311 surrogate and, thus, function continuously as an integral part of each CSM.

An example of first security component 106 of FIG. 1 is dongle 301 and an example of second security component 108 of FIG. 1 is dongle 303. Thus, in FIG. 3, MTU 300 and RTU 302 have secure communications over network 304 using modems 305 and 307 and the communication is secured by dongle 301, dongle 303, LKMC 311, and by KMC 310 as needed.

FIG. 3 also shows that a system for securing network communications scales up for multiple task-oriented components and security components. Of course, there are many different ways to arrange these components. In this example, multiple MTUs communicate with multiple RTUs over multiple networks. This communication is secured by multiple dongles in communication with LKMC 311.

Over network 304, MTU 300 has secure communications with RTU 302 through RTU 312. Over network 324, MTU 300 has secure communications with RTU 322 and other RTUs. Over network 334, MTU 300 has secure communications with RTU 332 and other RTUs.

MTU 300 has secure communications with RTU 302 over a communication path from MTU 300 to dongle 301 to modem 305 to network 304 to modem 307 to dongle 303 to RTU 302. Note that dongle 301 is interposed between MTU 300 and modem 305 and that dongle 303 is interposed between RTU 302 and modem 307. A communication path from MTU 300 to RTU 312 is from MTU 300 to dongle 301 to modem 305 to network 304 to modem 317 to dongle 313 to RTU 312.

MTU 300 has secure communications with RTU 322 over a communication path from MTU 300 to dongle 321 to modem 325 to network 324 to modem 327 to dongle 323 to RTU 322.

MTU 300 has secure communications with RTU 332 over a communication path from MTU 300 to dongle 331 to modem 335 to network 334 to modem 337 to dongle 333 to RTU 332.

Similarly, MTU 340 through MTU 370 have secure communications with various RTUs over various communication paths. MTU 340 has access to RTU 302 and RTU 312 through dongle 341 and modem 345. MTU 340 has access to RTU 322 through dongle 351 and modem 355. MTU 340 has access to RTU 332 through dongle 361 and modem 365.

While FIG. 3 shows an example configuration, many other configurations are possible. Some examples are:

1a. Many MTUs connect collectively to a single MTU dongle; or

1b. Many MTUs connect each to its own MTU dongle, which connect collectively to a single MTU modem; or

1c. Many MTUs connect each to its own MTU dongle and MTU modem, which latter connect collectively to a single network; and

2a. Many RTU modems with RTU dongles are connected to a common network representing one-to-many links; or

2b. Other networks have only a single RTU modem and RTU dongle, representing one-to-one links; and

3a. A single RTU connects to a single local RTU dongle; or

3b. Many RTUs connects to a single local RTU dongle.

FIG. 4 shows another example of a system for securing network communications. An MTU 400 has secured communications with its RTUs, RTU 402 through RTU 404, via a network 406. FIG. 4 shows a specific implementation of dongles as CSM and CSS dongles. MTU 400 is in communication with CSM dongle 408, which is in communication with both KMC 410 and modem 412. Modem 412 is in communication with modems 414 and 416. Modem 414 is in communication with CSS dongle 418, which is in communication with RTU 402, while modem 416 is in communication with CSS dongle 420 that is in communication with RTU 404. A CSM dongle is a not quite so small device interposed between an MTU and its directly connected master modem(s), which acts as a CSM. A CSS dongle is a small device interposed between a slave modem and its directly-connected slave RTU(s) which acts as a CSS. FIG. 4 shows an example of master-slave networking, but peer-to-peer networking and other kinds of networking also work.

There is a means of adding communications security to existing and future control systems, such as power transmission and distribution systems, oil and gas pipelines, and regional or municipal water and sewage management systems. Some embodiments also provide a basis for adding compatible communications security to internal local area networks (LANs) of process control systems, such as PlantScape® and Experion PKS™, which are available from Honeywell International Inc., Morristown, NJ.

There is hardware and software for retrofit situations and for central control of communications security and software products for new equipment or where upgrade of existing product software is the chosen course.

Users include control system operators worldwide who have a need to secure their communications systems and defend them against cyberattack. The present invention is exportable to all the countries in the world, subject to any government-imposed restrictions.

Communications between a control site and its distributed RTUs are secured in a control system, such as a SCADA system. A control system that is an industrial measurement and control system comprises:

1. A central host or master (a/k/a MTU), which may be redundant;
2. One or more field data gathering and control units or remotes (a/k/a RTUs);
3. A multi-point communications channel (or a collection of point-to-point communications channels, or a combination thereof) from the MTU(s) to the RTUs and from each RTU to the MTU(s); and

4. A collection of standard and/or custom hardware and software used to monitor and control remotely located field equipment.

Most SCADA systems exhibit predominantly open-loop control characteristics and use predominantly long distance communications, although some elements of closed-loop control and/or short distance communications are also used. Other types of control systems have predominantly closed-loop control characteristics. Still other types use predominantly short- or medium-distance communications or both. There is a wide variety of mixtures of such features in control systems.

Communications security (ComSec) is retrofitted to existing SCADA wide area networks (WANs) or is included directly in new SCADA equipment and networks. Communications security (ComSec) is defined as measures and control taken to deny any unauthorized person information derived from telecommunications and to ensure the authenticity of such telecommunications. Communications security includes cryptosecurity, transmission security, emission security, and physical security of ComSec material. Cryptosecurity is the component of communications security that results from the provision of technically sound cryptosystems and their proper use. When the existing equipment needs to remain unmodified, one approach is to place cyberprotective devices on the ends of the links at a point of exposed connection between the communicating end equipment and the intermediary modems that provide the network's physical signaling. For older equipment and systems, such exposed connection points usually exist, typically taking the form of RS-232 cables and connectors between equipment and nearby modems. Some example embodiments include the following:

1. A small connectorized package known as a dongle, the CSS dongle, at each field site of the network, which is interposed between a 9-pin RS-232/RS-423 serial port of a modem and its attached RTUs.

2. A somewhat larger dongle, the CSM dongle, at the central control site of the network that is interposed between a 9-pin RS-232/RS-423 serial port of an MTU and its attached modem(s).

3. A very small dongle, the power dongle, that can be plugged in series with the CSS dongle to power the CSS dongle when its available parasitically-derived power is insufficient.

4. The smaller dongle's software that is capable of being incorporated into an RTU by the RTU software vendor.

5. A PCI card form of the larger dongle, the CSM PCI card, that is interposed logically and perhaps physically in the information flow between the MTU and its attached modem(s).

6. Variants of (1), (2), and (5) above, supporting other types of serial ports, such as 8-pin and 25-pin RS-232/RS-423 connectors, 37-pin RS-422 connectors, and the like.

7. Variants of (2) above where the MTU connection is via USB, firewire, or a similar serial bus.

8. Variants of (3) above, supporting non-SCADA instrumentation, such as field instruments on an appropriate fieldbus.

9. Variants of (5) above without serial ports that provides ComSec and a dual high speed Ethernet connection for time-critical process control LANs. For most utility, high-resolution time synchronization is also included.

The larger CSM dongle, (2) above, and some of the unplanned variants of the smaller CSS dongle are expected to need an external low-voltage power source. The CSS dongle, (1) above, is powered parasitically from its RS-232/RS-423 interfaces to a local modem and local equipment, such as an RTU.

The ComSec dongles and the power dongle target modems that are connected to an MTU or to one or more RTUs by an RS-232/RS-423 serial cable and connectors. The CSS software targets RTU vendors, whose RTUs include the following features:

1. Non-volatile rewritable program and data storage of at least 8 kB that are rewritable at least 20 times, e.g., flash memory.

2. Non-volatile rewritable data storage of at least (M+2)x64 B that can be rewritten at least 10,000 times, e.g., EEPROM, where M is the number of distinct multicast groups to which the device belongs.

The CSM PCI card targets MTU vendors whose equipment has an available PCI slot and which sometimes needs support for multiple concurrent RTU communications subnetworks.

For CSS and CSM dongles, there is no inherent restriction on the locale of manufacture of any hardware embodiment, because preferably no confidential or government restricted (for example, export controlled) software or hardware is present in either the embodiment or the manufacturing process at time of manufacture. There is a method for product preparation for distribution and sale. After manufacture and before placement into the distribution chain, a CSS or CSM dongle is sent to a trusted third party to preconfigure it with software and precommission it with unique identifying information and cryptographic secrets. A trusted third party installer is an agent that installs initial ComSec software and device-unique information into newly manufactured hardware devices before they are inserted into product distribution channels. This information is retained for escrow at a secure facility for use in assisting the system owner in failure recovery and for law enforcement use under a recognized court order. There are many reasons to use a trusted third party. First, it ensures that only the intended software is loaded into the device, so that the device may be manufactured in untrusted countries and facilities by uncleared personnel. Second, it supports revenue and customer service goals. Finally, it ensures compliance with government mandated requirements on the content of the software or the escrow of keys.

A trusted third party powers up one or more devices of a common type and downloads in parallel to their flash memories:

1. A boot loader that deciphers stream-enciphered download images given the appropriate key;
2. A download traffic encryption key (TEK); and
3. The current version of the software appropriate to the device, stream enciphered under that TEK.

It then downloads to each device separately:

1. A unique device class identifier (ID) and serial number;
2. A unique key for the device, known as the birth key encryption key (KEK); and
3. One or more encrypted versions of that birth KEK, where each encryption key is either a symmetric or public key common across all CSMs and CSSs.

Enciphering and deciphering involve ciphers. A cipher is a cryptographic system in which units of plaintext (unencrypted information) data are substituted according to a predetermined key, resulting in ciphertext (encrypted information) data. There are different kinds of ciphers, for example block ciphers. A block cipher is a type of symmetric cipher that transforms a fixed-length block of plaintext into a block of ciphertext data. This transformation takes place under the action of a user-provided secret key. Applying the reverse transformation to the ciphertext block using the same secret key deciphers the block, resulting in the original plaintext. The fixed length is called the block size, which for modern block ciphers is typically 128 bits. Ciphertext is enciphered information. Plaintext is unencrypted information. Cleartext is synonymous with plaintext. To encipher is to convert plaintext into an unintelligible form by means of a cipher. A symmetric cipher is a reversible cipher which uses the same key to transform a plaintext data stream into a ciphertext data stream, or vice versa, depending on the direction of operation. A symmetric stream cipher is any symmetric cipher that changes how it behaves during a message. Such ciphers can be designed to be exceptionally fast, much faster than any block cipher. They usually work on small units of text, generating a keystream that is combined reversibly with the text to transform plaintext to ciphertext and vice versa, depending on the direction of operation.

In some embodiments, the one public key is known to all CSMs, perhaps by preconfigured code, and another public key is known for use in key recovery assistance as ordered by competent legal authority. The preconfigured and precommissioned devices are then repackaged, after which they are ready for distribution and sale.

A public key is the unprotected key of public key cryptography, used for encryption and validating digital signatures. A private key is the protected key of public key cryptography, used for decryption and digital signing. Public key cryptography is the type of cryptography in which the encryption process is publicly available and unprotected, but in which a part of the decryption key (the private key) is protected so that only a party with knowledge of both parts of the decryption process can decrypt the ciphertext. A key encryption key (KEK) is a cipher key used to encrypt other keys. A traffic encryption key (TEK) is a symmetric cipher key used to encrypt plaintext and decrypt ciphertext or to super-encrypt and super-decrypt ciphertext.

There are installation and update methods. A control system operator has one or more CSM devices and an initial batch of CSS dongles or RTUs containing CSS software. Some control system operators have one CSM per MTU and one CSS per RTU modem or per RTU where a modem is multidropped to many RTUs, plus an adequate number of spares of each.

There is a method for establishing a ComSec system. Each CSM is capable of establishing its own unique and intentionally non-interoperable ComSec system. This establishment occurs when an agent of the end user configures the CSM. Subsequent CSM and CSS devices are made members of the same ComSec system by any CSM that is currently a member of the system, which initially is just the first configured CSM.

There is a method for configuring and commissioning the initial CSM. The user agent that configures and commissions a CSM dongle applies power to the dongle and establishes a management dialogue with the dongle through the dongle's Ethernet port.

Through the management dialogue, the user agent specifies the communications protocol used by the control system. This specification is in the form of a selection among listed alternatives or in the form of a very small file, which describes the communications protocol to be secured, which is transferred to the CSM.

The user agent specifies the method by which the user's operational ComSec agents will authenticate commands to the ComSec system once it is operational, which occurs immediately after the CSM has been configured and commissioned. A common method would be the specification of two distinct pieces of information that are provided either by one or two individuals. This is known as two-factor authentication. More complex authentication through weighted secret sharing is supported.

The user agent specifies the parameters of the key escrow provided by the system, such as the need for and duration of key escrow, the set of

Internet or intranet network addresses to which escrowed keys should be sent, which may be a null set, and the desired immediacy or frequency of this transmission of escrowed keys to the specified address.

At this point, the CSM has been configured and commissioned and is prepared to form its own isolated ComSec system. The CSM generates the following items:

1. A unique system ID comprising its own device serial number concatenated with a count of the number of times it has created such a system ID.
2. A new key called the system KEK.
3. A unique system device ID, for example, an ID formed from the system ID concatenated with the count of the number CSMs which this CSM has commissioned, which is one (itself).
4. A second new key called a personal KEK.

At this point, the CSM has established its own isolated ComSec system.

FIG. 5 shows a method for configuring and commissioning security components, such as additional CSMs according to the present invention. An additional CSM is included in the ComSec system established by the first CSM.

A user agent authorizes the initial CSM, or any other CSM already included in the ComSec system of the initial CSM, to activate its commissioning functions. This authorization is authenticated according to the policy established when the ComSec system was formed by the first CSM, or as subsequently modified. Such authorization of the commissioning port expires after a predetermined period of non-use of the commissioning functions, typically after about 5 to 10 minutes of non-use.

A user agent couples the distributed computing environment (DCE) port of the new CSM to the commissioning port of the CSM whose commissioning functions have been authorized, while that authorization is still in force. The commissioning CSM configures the new CSM with all of the information established by the user agent so that the new CSM has the same configuration and policy parameters as the commissioning CSM.

The commissioning CSM requests the new CSM's birth KEK, as encrypted under the first of the system-wide keys. The commissioning CSM decrypts that information and then uses the birth KEK of the new device to establish a session key (a/k/a traffic encryption key (TEK)) for the remaining information exchanges during the commissioning process.

The commissioning CSM generates: a unique system device identifier (ID) and a new key for the new CSM. For example, the unique system device ID is based on its own system device ID concatenated with a count of the number of CSMs which this CSM has commissioned, which is at least one (the new CSM). The new key is called the new CSM's personal KEK. The commissioning CSM transfers the unique system device ID and the new key to the new CSM using the just-established TEK.

At this point, the new CSM has been made part of the commissioning CSM's ComSec system and is ready for deployment in the SCADA control center, a secondary control center, or wherever else the SCADA system operator desires. To effect that deployment, the new CSM is connected by its Ethernet port to the other CSMs of the same ComSec system. Such connection can be local or through the Internet or through a WAN or other communications media. The CSMs use this connection to share keys and other information about discovered CSMs and CSSs. The CSM-induced traffic on this connection is extremely low.

If it is to provide operational communications protection for an existing MTU, the CSM is also installed in or connected to that MTU so that the CSM is able to encrypt and decrypt the MTU's communications with its RTUs.

Note that CSMs have other uses. For example, some CSMs are intended as commissioned spare parts or for use in offsite commissioning of other devices and need not function with an MTU.

FIG. 6 is a block diagram of a method for deploying a security component, such as a dongle CSM according to the present invention. The person deploying a dongle CSM installs the dongle on the MTU's connection to its modem(s) by momentarily interrupting the MTU's connection to its modem(s) to insert the dongle between the MTU's RS-232/RS-449 connector and the attached serial cable. Power is applied to the CSM dongle either before or after it is inserted. The dongle begins to function almost transparently, observing but not modifying the SCADA communications. However, it does introduce an additional delay (typically of one-character for both inbound and outbound messaging) into the SCADA system's scan cycle due to its message character serialization and deserialization processes. Note that this delay is reducible to one bit on low-speed networks through more aggressive CSM dongle software and hardware design.

There are various methods of operation. One method of operation is for adding ComSec to the control system communications. One method of operation for adding ComSec to the control system communications is a method for discovery of unicast RTU addresses. While operating almost transparently, the CSM analyzes the message headers of the messages it forwards, isolating the unicast addresses and multicast addresses in use on the network. It retains these addresses to manage its CSSs.

Periodically during its operation, the CSM delays giving its attached MTU a clear-to-send signal, forcing the MTU to wait while the CSM communicates with some RTU's CSS on its own. The length of this delay is short, perhaps 50 ms on a 2400 bit/s communications network, and proportionately less at higher data rates. During this interval, the CSM sends a ComSec poll message to one of the RTU unicast addresses that the CSM has observed and saved, and which is not known to have an associated CSS. The form of the ComSec poll is protocol specific, but it is always a message that will be ignored or treated as an error by an RTU that does not have an interposed CSS.

If there is a newly installed CSS at the polled address, the CSS responds to the CSM with a secure ComSec reply message giving the CSS's system ID and the list of unicast addresses to which the CSS's RTUs have responded, all authenticated with the KEK the CSM wrote into the CSS. The CSM associates the CSS's ID with the polled address, and with any other addresses that the CSS has given in its response. The CSM stops further polling of those addresses unless the CSS and its RTUs should become nonresponsive.

Another method of operation is a method for establishing ComSec for discovered addresses. At a time of its choosing, the CSM sends the CSS a new session key, stream enciphered under the CSS's KEK, and associates that key with the unicast RTU address(es) of the CSS. A session key is a TEK for the set of messages that comprise a communications session. From that point on, all communications with the CSS and its RTU(s) are stream-enciphered and secured, unless the CSS becomes nonresponsive or is replaced by another dongle, in which case the low-frequency poll of the affected address is restarted.

If there are multiple CSMs for redundant MTUs, the CSM shares: the CSS system ID, the newly-created session key, and the set of addresses associated with that session key with its peer CSMs via their shared Ethernet connection. This sharing has sequence numbers, so after powerup, each CSM can inquire of the others whether any update messages have been lost, and if so request a replacement copy of either the lost information or the full database.

These tables of CSS system IDs, keys, and set of associated addresses are retained in memory, such as the internal RAM of the CSM. If an implementation has CSM hardware with the EEPROM external to the microcontroller chip, then at least the keys are also written in enciphered form to a memory, such as key storage EEPROM within the CSM, under a key created by the CSM for that purpose, after copying any prior key information for that CSS from the EEPROM to a large key escrow flash memory within the CSM. EEPROM is non-volatile memory which has been specially constructed to be erasable and capable of being rewritten a large number of times, typically 10^6 times. Flash memory is non-volatile memory, of higher density and lower cost per bit than EEPROM, which has been specially constructed to be erasable and capable of being rewritten a limited number of times, typically 50-10,000 times. Thus, in one embodiment, operational key information is stored within the CSM's RAM, while an enciphered form is retained in the non-volatile key storage EEPROM and prior keys are retained in enciphered form in the non-volatile key escrow flash memory when key escrow is configured.

Another method of operation is a method for establishing ComSec for some multicast addresses before full system ComSec has been established. Multicast addresses other than the broadcast address are discovered in messages from the MTU, but the set of RTUs that is addressed by such a multicast address is usually not discoverable. Unlike the recipients of unicast messages, multicast message recipients do not generate an immediate reply message from which their identity can be learned. Thus, the CSM assumes the entire set of CSSs are potential intended recipients of each multicast address, except when explicit information on set membership is provided through an extension of CSM configuration.

For each distinct multicast set, as soon as all of the RTU addresses in that set are known to have interposed CSSs, and those CSSs have been given the key(s) for the multicast address(es) associated with that set, then the CSM notifies the involved CSSs that it will now apply ComSec protection to messages addressed to multicast addresses of that set. Thus, the CSM provides ComSec protection for all network addresses, including any multicast address(es), as soon as all of the RTUs in the network have interposed CSSs and the appropriate session keys are shared.

If incremental protection of multicast groups is desired before CSSs have been interposed at all RTUs, then the CSM needs outside assistance before it can secure those groups while leaving other groups unsecured. Because the CSM cannot infer the membership of these multicast groups on its own, it learns the information from the control system operator.

During normal operation, even while operating completely transparently, the CSM observes the multicast addresses in messages that it is sending. It accumulates this list and provides it on request to the control system operator via a network, such as an Ethernet connection.

Whether in a delayed response, or on his/her own, an agent of the system operator sends a list of the set of RTU unicast addresses that are members of each multicast set to the CSM. Upon receipt of the list, the CSM analyses the multicast group membership as previously described, creates new keys as appropriate, and sends messages to each of the affected

CSSs, giving them the appropriate subset of the new keys and the multicast group address(es) associated with each of those keys.

Another method of operation is a method for ComSec overlay of control system communications. This method includes how ComSec is applied to and modifies the RTU messaging. With respect to the pre-ComSec communications, the CSM and CSSs have the following goals: (1) add ComSec to some or all of the messaging on the WAN, (2) minimize the delay they induce in the control system communications cycle, and (3) minimize the impact of this addition on the RTUs and the MTU(s).

Another method of operation is a method for replacement of a CSM in an operational ComSec system. A replacement CSM is added to an operational ComSec system in the same manner as a redundant CSM. The CSM is configured and commissioned appropriately for the system. The CSM is connected to the network of operational CSMs via their Ethernet ports so that the operational CSMs are capable of sharing session keys and related information with the new CSM. At this point, the new CSM is added as another (redundant) CSM or it replaces any of the existing CSMs.

If there is no operational CSM from which the system's operational CSM database is able to be retrieved, then the replacement CSM is installed in the system just as the first CSM when that CSM was commissioned. On powerup, a commissioned CSM stops transmissions from the associated MTU by dropping its DTR and CTS signals to the MTU. Also on powerup, the commissioned CSM sends (preferably to a broadcast address) a protocol-specific distinguished cleartext message sequence to all attached CSSs to notify them of the need to turn off temporarily all ComSec for CSM/CSS communications (but not for MTU/RTU communications).
Another approach is to use symmetric key broadcast authentication. A single precommissioned backup CSM sometimes detects in run-time the

specific system (of the same protocol type) of a common system owner that is to be recovered. Reversion to cleartext MTU/RTU transmission is not permitted to forestall pseudo-MTU-recovery attacks on the SCADA system. Reversion to cleartext for the CSM/CSS communications is just temporary to limit the extent of potential attacks on the ComSec control communications.

As rapidly as possible, the CSM queries each protocol-determined potential unicast RTU address for the enciphered version of any associated CSS's backup KEK, which is sent as a cleartext reply to the CSM. The backup KEK was enciphered under the CSM system KEK and given to the CSS as part of the CSS commissioning.

Upon receipt of the enciphered backup KEK, the CSM deciphers it and determines whether this is a new or a previously discovered CSS. If it is a new CSS, then the CSM generates a new primary KEK for the CSS, enciphers it under the just-recovered CSS backup KEK, and sends that new KEK to the CSS.

When it receives the new KEK, the CSS returns its list of associated RTU addresses, including any multicast addresses that have associated session keys, to the CSM in one or more messages protected by the new KEK. The CSM uses this information to regenerates the failed CSM's database of RTU addresses and common address sets.

As the CSM receives this information, it generates session keys for the unicast address(es) of each responding CSS, enciphered under the CSS's new KEK, and sends those session keys to the CSS.

When the entire database has been reconstituted, the CSM determines the non-overlapping multicast session sets and then generates a new session key for each set of multicast addresses. Each CSS that participates in multicast sessions, including the broadcast session, is sent the set of session keys associated with the multicast (including broadcast) group address(es) that its RTUs recognize, enciphered under that CSS's KEK. At this point, ComSec has been restored to the SCADA system.

It is to be understood that the above description is intended to be illustrative and not restrictive. Many other embodiments will be apparent to those of skill in the art upon reviewing the above description, such as adaptations of the present invention to various hardware, software, and firmware forms. Various types of networks, such as local area networks are contemplated by the present invention, even though some minor elements would need to change to better support the low-delay, peer-to-peer environment common to such networks. The present invention has applicability to fields outside SCADA networks, such as field instrument networks, communications networks of distributed control system (DCS), enterprise building integrator (EBI) systems, and other time-critical systems. Therefore, the scope of the present invention should be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. A method for configuring and commissioning, comprising:

specifying a communications protocol for a first security component;

specifying an authentication method for said first security component; and

specifying key escrow parameters for said first security component.

2. The method according to claim 1, further comprising:

authorizing said first security component to activate a commissioning method;

while authorization is in force, coupling a second security component to a port of said first security component; and

configuring said second security component by said first security component.

3. The method according to claim 2, wherein configuring said second security component by said first security component comprises:

requesting a first key encryption key (KEK) from said second security component;

decrypting said first KEK to establish a session key;

generating, by said first security component, an identifier for said second security component;

generating, by said first security component, a second key encryption key (KEK); and

sending said identifier and said second KEK to said second security component using said session key.

4. The method according to claim 3, wherein said identifier is a unique system device identifier.

5. The method according to claim 1, wherein said first security component is a new security component and further comprising:

connecting said first security component to a network so that other security components share a session key with said first security component.

6. The method according to claim 5, wherein said first security component is a ComSec master (CSM), a CSM database is either not operational or not retrievable, said second security component is a ComSec slave (CSS), and said other security components are other ComSec slaves (CSSs).

7. The method according to claim 6, further comprising:

if said CSS is newly discovered by said CSM, generating, by said CSM, a third key encryption key (KEK) for said CSS, enciphering said third KEK under said second KEK, and sending said third KEK to said CSS;

receiving, by said CSM, at least one list in at least one message protected by said third KEK from said CSS;

using, by said CSM, said at least one list to regenerate at least a portion of said CSM database; and

generating, by said CSM, new session keys based on said at least one list and sending said new session keys to said CSS and said other CSSs such that each CSS receives only session keys needed for communications with at least one associated remote terminal unit (RTU).

8. The method according to claim 7, further comprising:

upon powerup, stopping, by said CSM, communication from a master terminal unit (MTU); and

upon powerup, stopping, by said CSM, communication from said CSS to said CSM.

9. The method according to claim 7, further comprising:

requesting, by said CSM from a remote terminal unit (RTU), an address; and

deciphering, by said CSM, said address to determine whether said CSS is new.

10. The method according to claim 7, wherein said list is associated RTU addresses, including any multicast addresses with associated session keys.

11. The method according to claim 9, wherein said address is a unicast address for an enciphered version of a backup key encryption key (KEK) for said CSS.

12. A method for deploying, comprising:

installing a security component on a connection of a task-oriented device to a modem by interrupting said connection to insert said security component between said task-oriented device and said modem; and

applying power to said security component;

wherein said security component alters a communication to or from said task-oriented device.

13. The method according to claim 12, further comprising:

configuring and commissioning said security component, before installing.

14. The method according to claim 13, wherein said task-oriented device is a master terminal unit (MTU) and said security component is a ComSec master (CSM).

15. A system for configuring and commissioning, comprising:

a first security component to configure a second security component so that said second security component has a substantially similar configuration as said first security component; and

an agent to authorize said first security component to activate said first security component's commissioning function and to couple said second security component to a port of said first security component, while authorization is in force.

16. The system according to claim 15, wherein said first security component is a commissioning ComSec master (CSM) and said second security component is a new CSM.

17. A computer-readable medium having computer-executable instructions for performing a method, comprising:

specifying a communications protocol for a first security component;

specifying an authentication method for said first security component; and

specifying key escrow parameters for said first security component.

18. The computer-readable medium according to claim 1, further comprising:

authorizing said first security component to activate a commissioning method;

while authorization is in force, coupling a second security component to a port of said first security component; and

configuring said second security component by said first security component.

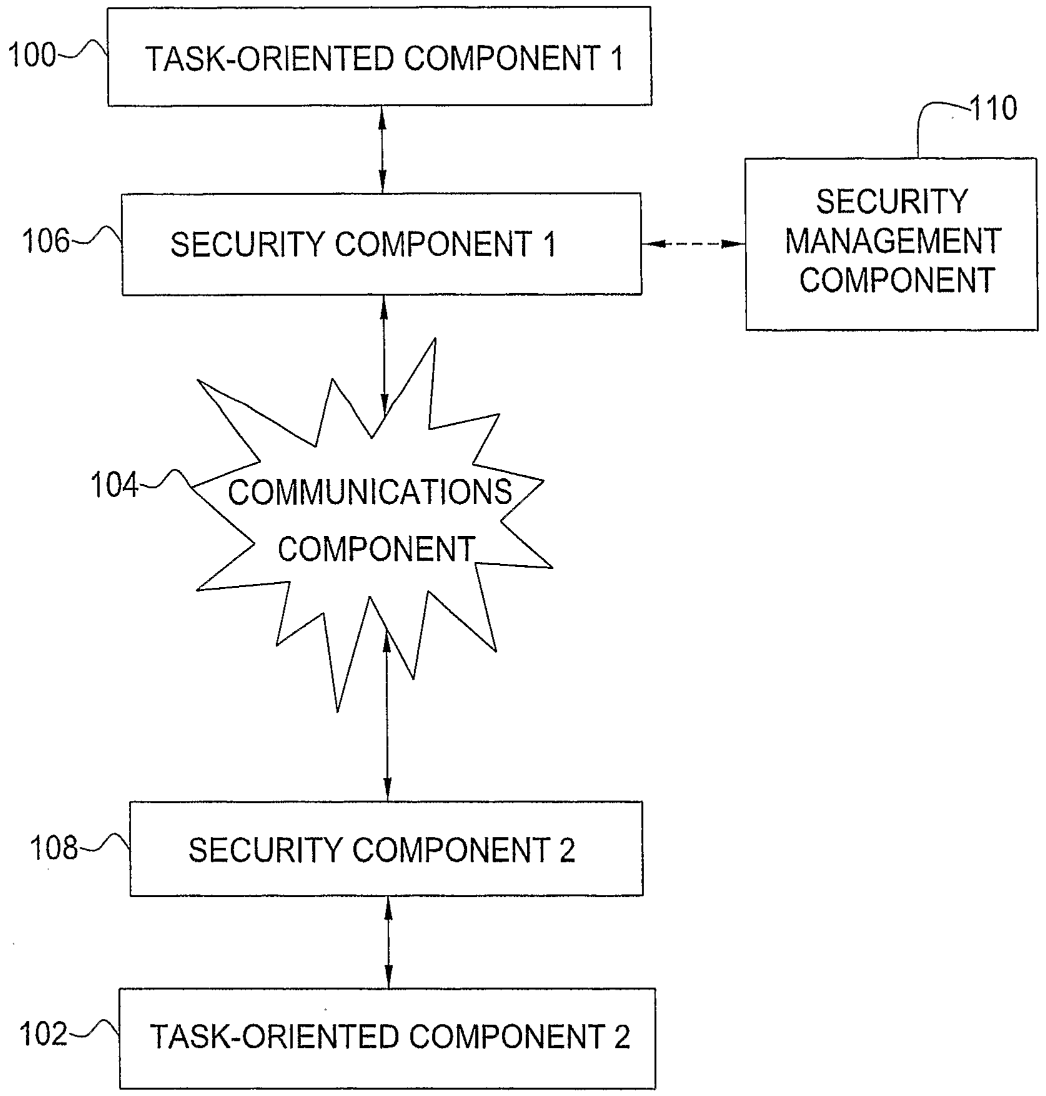

100
TASK-ORIENTED COMPONENT 1
110
106
SECURITY COMPONENT 1
SECURITY MANAGEMENT COMPONENT
104
COMMUNICATIONS COMPONENT
108
SECURITY COMPONENT 2
102
TASK-ORIENTED COMPONENT 2

FIG. 1

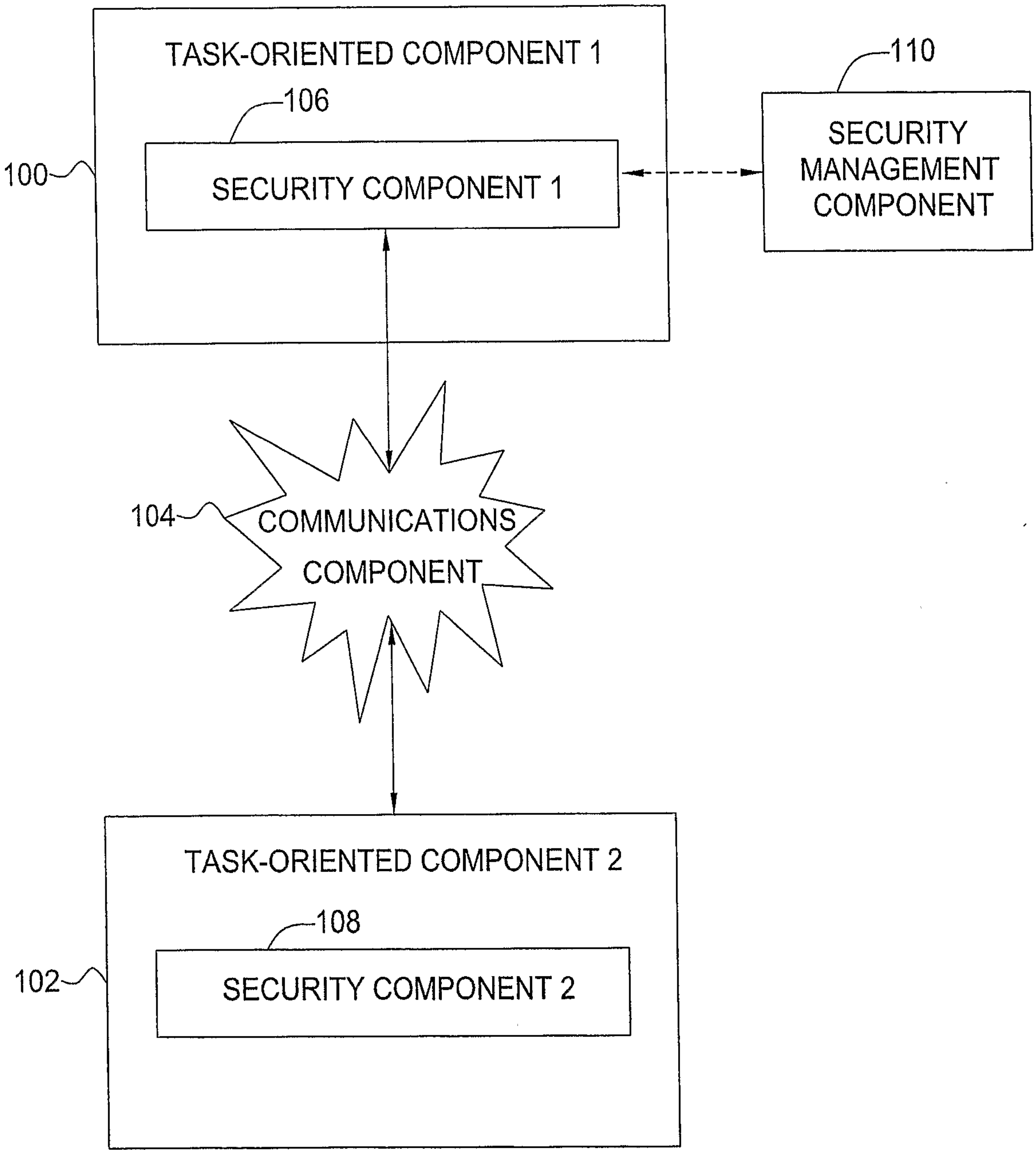

TASK-ORIENTED COMPONENT 1
106
SECURITY COMPONENT 1
100
110
SECURITY MANAGEMENT COMPONENT
104
COMMUNICATIONS COMPONENT
TASK-ORIENTED COMPONENT 2
108
SECURITY COMPONENT 2
102

FIG. 2

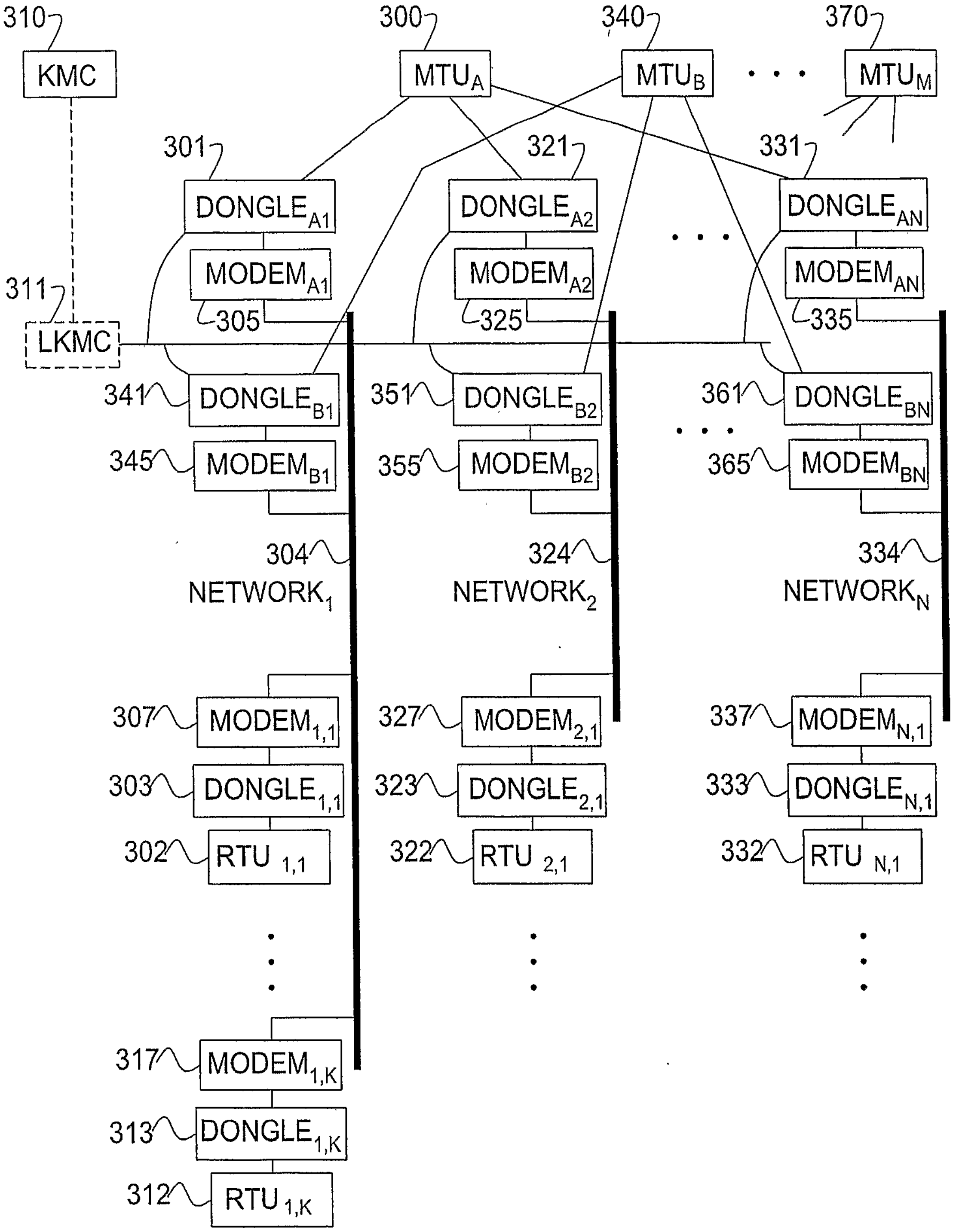

310
KMC
300
MTU$_A$
340
MTU$_B$
370
MTU$_M$
301
DONGLE$_{A1}$
MODEM$_{A1}$
305
321
DONGLE$_{A2}$
MODEM$_{A2}$
325
331
DONGLE$_{AN}$
MODEM$_{AN}$
335
311
LKMC
341
DONGLE$_{B1}$
345
MODEM$_{B1}$
351
DONGLE$_{B2}$
355
MODEM$_{B2}$
361
DONGLE$_{BN}$
365
MODEM$_{BN}$
304
NETWORK$_1$
324
NETWORK$_2$
334
NETWORK$_N$
307
MODEM$_{1,1}$
303
DONGLE$_{1,1}$
302
RTU $_{1,1}$
327
MODEM$_{2,1}$
323
DONGLE$_{2,1}$
322
RTU $_{2,1}$
337
MODEM$_{N,1}$
333
DONGLE$_{N,1}$
332
RTU $_{N,1}$
317
MODEM$_{1,K}$
313
DONGLE$_{1,K}$
312
RTU $_{1,K}$

FIG. 3

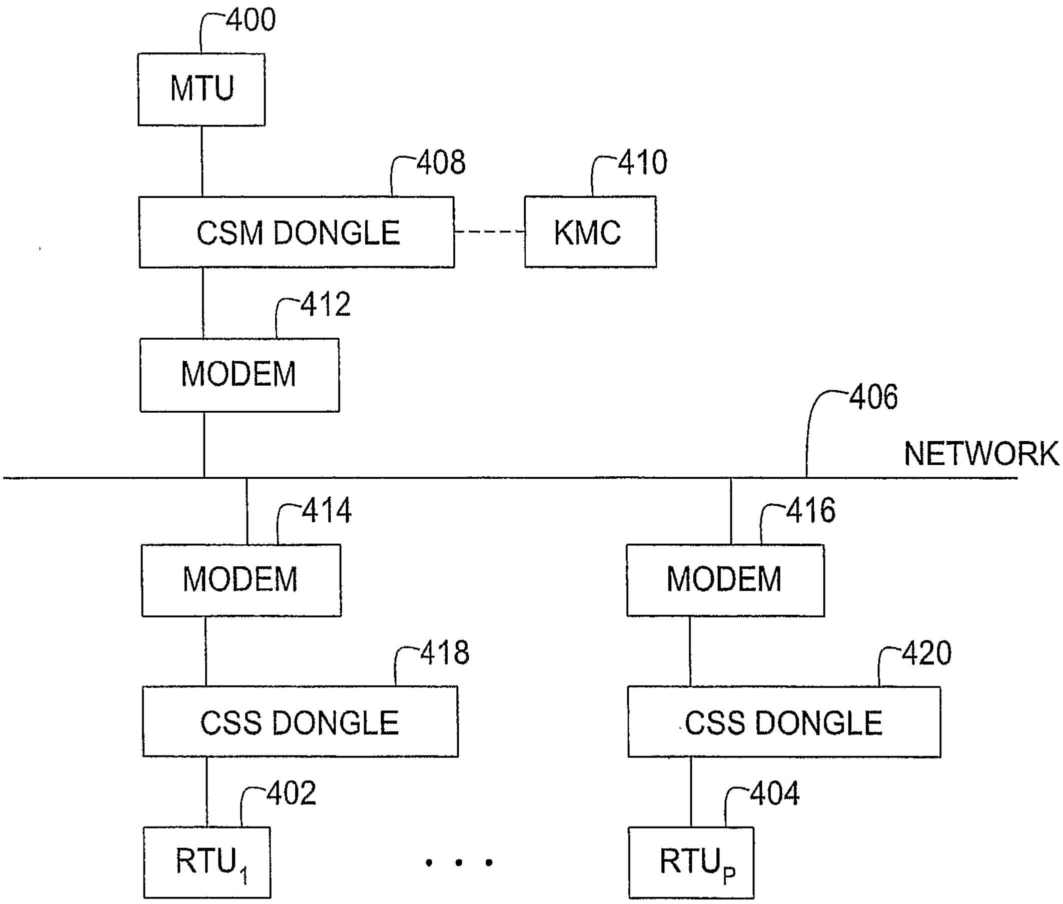

400
MTU
408
CSM DONGLE
410
KMC
412
MODEM
406
NETWORK
414
MODEM
416
MODEM
418
CSS DONGLE
420
CSS DONGLE
402
RTU_1
. . .
404
RTU_P

FIG. 4

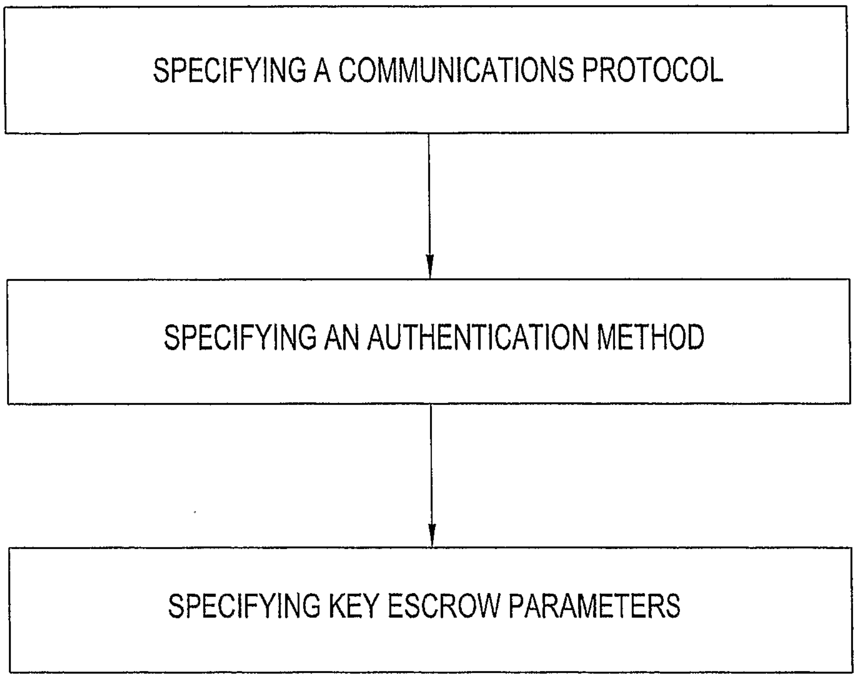

SPECIFYING A COMMUNICATIONS PROTOCOL
SPECIFYING AN AUTHENTICATION METHOD
SPECIFYING KEY ESCROW PARAMETERS

FIG. 5

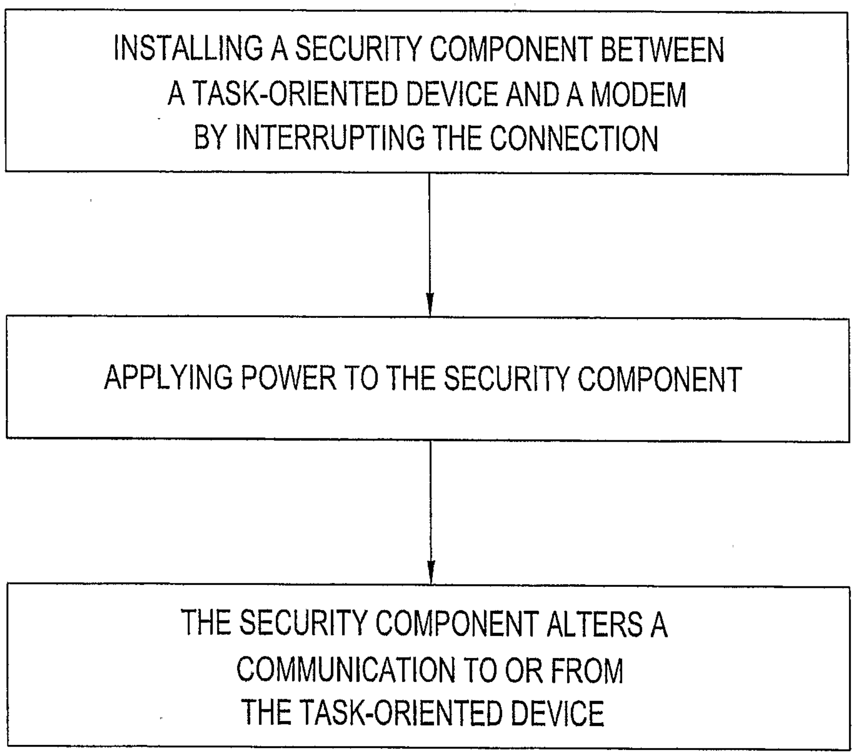

INSTALLING A SECURITY COMPONENT BETWEEN A TASK-ORIENTED DEVICE AND A MODEM BY INTERRUPTING THE CONNECTION
APPLYING POWER TO THE SECURITY COMPONENT
THE SECURITY COMPONENT ALTERS A COMMUNICATION TO OR FROM THE TASK-ORIENTED DEVICE

FIG. 6